

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 5, 2006, 09:47:58 ; Search time 107 Seconds  
(without alignments)  
3473.990 Million cell updates/sec

Title: US-10-509-950-1

Perfect score: 4235  
Sequence: 1 MSEGQDNLQALAEAGTEGFE.....SASCRTNCASMSNMGWQYK 813

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

A-Geneseq2.8:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1980s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2000s:\*
- 5: geneseqp2000s:\*
- 6: geneseqp2000s:\*
- 7: geneseqp2000s:\*
- 8: geneseqp2000s:\*
- 9: geneseqp2000s:\*
- 10: geneseqp2000s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | ID | Description        |
|------------|--------|-------------|--------|----|--------------------|
| 1          | 4235   | 100.0       | 813    | 7  | ADDB9836 Human CAM |
| 2          | 4235   | 100.0       | 813    | 7  | ADG17598 Human Brl |
| 3          | 4208   | 99.4        | 847    | 7  | ADG17594 Human TAR |
| 4          | 4111   | 97.1        | 793    | 7  | ADG17600 Human Brl |
| 5          | 4084   | 96.4        | 827    | 7  | ADG17596 Human Brl |
| 6          | 3586   | 84.7        | 682    | 5  | ABP43676 Human TAR |
| 7          | 3421   | 80.8        | 807    | 7  | ADG17603 Human TAR |
| 8          | 3414.5 | 80.6        | 651    | 7  | ADG17605 Human TAR |
| 9          | 3398.5 | 80.2        | 648    | 7  | ADG17604 Human TAR |
| 10         | 2513.5 | 59.4        | 645    | 4  | ABG04987 Human pro |
| 11         | 2047   | 48.3        | 438    | 4  | ABG02090 Novel hum |
| 12         | 1737.5 | 41.0        | 1044   | 8  | ADG09546 Human pro |
| 13         | 1510   | 35.7        | 971    | 7  | ADG56304 Human pro |
| 14         | 1493   | 35.3        | 976    | 6  | ABR58323 Human pro |
| 15         | 1455   | 34.4        | 1027   | 6  | ABR43214 Human pro |
| 16         | 1451   | 34.3        | 958    | 6  | ABR58322 Human pro |
| 17         | 1420.5 | 33.5        | 433    | 4  | ABG02088 Novel hum |
| 18         | 1332.5 | 31.5        | 840    | 7  | ADG37199 Nuclear f |
| 19         | 1294.5 | 30.6        | 854    | 7  | ADG37201 Nuclear f |
| 20         | 1020   | 24.1        | 366    | 4  | ABG04986 Novel hum |
| 21         | 944    | 22.3        | 371    | 4  | ABG02089 Novel hum |
| 22         | 760.5  | 18.0        | 723    | 7  | ADG37197 Nuclear f |
| 23         | 582    | 13.7        | 123    | 4  | ABR15877 Human ner |

|    |       |      |      |   |                     |
|----|-------|------|------|---|---------------------|
| 24 | 491   | 11.6 | 162  | 4 | ABR17003 Human ner  |
| 25 | 473   | 11.2 | 107  | 7 | ADG55582 Human pro  |
| 26 | 430   | 10.2 | 89   | 5 | AAW47778 Human har  |
| 27 | 430   | 10.2 | 89   | 7 | ADG17606 Human pro  |
| 28 | 430   | 10.2 | 89   | 7 | ADG17602 Human Brl  |
| 29 | 423   | 10.0 | 89   | 3 | AAW01375 Human pro  |
| 30 | 410.5 | 9.7  | 1805 | 4 | ABG55262 Drosophila |
| 31 | 346.5 | 8.2  | 94   | 4 | ABG04984 Novel hum  |
| 32 | 346.5 | 8.2  | 94   | 4 | ABG02085 Novel hum  |
| 33 | 344.5 | 8.1  | 88   | 7 | ADG17607 Human pro  |
| 34 | 311.5 | 7.4  | 78   | 5 | ABR33783 Human ORF  |
| 35 | 308   | 7.3  | 79   | 4 | ABG04985 Novel hum  |
| 36 | 308   | 7.3  | 79   | 4 | ABG02086 Novel hum  |
| 37 | 293   | 6.9  | 2004 | 5 | ABG55113 Human tra  |
| 38 | 293   | 6.9  | 2004 | 9 | ADG07185 Human tra  |
| 39 | 293   | 6.9  | 2004 | 9 | ADG07185 Human tra  |
| 40 | 291.5 | 6.9  | 1400 | 9 | ADG85501 Human tra  |
| 41 | 291   | 6.9  | 2000 | 9 | ADY14779 PRO polyP  |
| 42 | 291   | 6.9  | 2703 | 4 | ABR63239 Drosophila |
| 43 | 286   | 6.8  | 1390 | 5 | ABG70019 Larval vi  |
| 44 | 286   | 6.8  | 1390 | 5 | ABG55152 Human NOV  |
| 45 | 286   | 6.8  | 1390 | 8 | ADN61955 Human nov  |

## ALIGNMENTS

RESULT 1  
ADDB9836  
ID ADDB9836 standard; protein, 813 AA.  
AC  
ADDB9836;  
DT 29-JAN-2004 (first entry)  
DE Human CAM-regulated phosphoprotein hTRAP.  
XX Human; hTRAP; TRAP; neurodegenerative disease; Alzheimer's disease;  
KW gene therapy; neuroprotective; noctropic.  
XX Homo sapiens.  
OS  
FH Key Location/Qualifiers  
FT Misc-difference 68 /note= "Encoded by AGY"  
XX  
XX W02003083482-A1.  
XX PD  
XX 09-OCT-2003.  
XX  
XX 01-APR-2003; 2003WO-EP003364.  
XX PR 02-APR-2002; 2002EP-00007522.  
XX PR 02-APR-2002; 2002US-0368970P.  
XX PA (EVOT-) EVOTEC NEUROSCIENCES GMBH.  
XX PI Hipfler R, Hanes J, Von Der Kammer H, Pohlner J;  
XX WPI; 2003-804100/75.  
XX DR N-PSDB; ADDB9837.  
XX  
XX PT New nucleic acid useful for preparing a composition for treating,  
XX PT preventing, diagnosing, prognosticating or monitoring the progression of  
XX PT a neurodegenerative disease, e.g. Alzheimer's disease.  
XX PS Claim 8; SEQ ID NO 1; 69pp; English.  
XX  
XX The present sequence is the protein sequence of human CAM-regulated  
XX phosphoprotein TRAP (or hTRAP), a previously undescribed human isoform  
XX of murine CAM-regulated phosphoprotein ARPP-21. The invention discloses  
XX the differential expression of hTRAP in specific brain regions of  
XX Alzheimer's disease patients. On the basis of this finding, a method is

CC provided or diagnosing or prognosticating a neurodegenerative disease,  
CC in particular Alzheimer's disease, or for determining whether a subject  
CC is at increased risk of developing such a disease. The gene encoding  
CC hTRAPP is used in therapeutic and prophylactic methods for treating or  
CC preventing Alzheimer's' disease and related neurodegenerative disorders.  
CC A method of screening for modulating agents of neurodegenerative diseases  
CC is also provided. Transgenic animals comprising a non-native gene  
CC sequence coding for hTRAPP are useful for screening, testing or  
CC validating compounds, agents or modulators, and developing diagnostics or  
CC therapeutics for treating neurodegenerative disease, particularly  
CC Alzheimer's disease. hTRAPP protein is useful as a diagnostic target for  
CC detecting, or screening target for reagents or compounds for preventing,  
CC treating or ameliorating a neurodegenerative disease, preferably  
CC Alzheimer's disease.

**SQ** Sequence 813 AA;

| Query Match | Score | DB | Length |
|-------------|-------|----|--------|
| 100.0%      | 4235  | 7  | 813    |

Best Local Similarity 100.0%; Pred. No. 3.5e-278;

```
Matches 813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

|    |  |     |
|----|--|-----|
| QY | MSBOEDLNQALAEBCGTEOETATPENGIYKSESLEDEEKLEIQRLEAONQERKKSQA        | 60  |
| Db | 1 MSBOEDLNQALAEBCGTEOETATPENGIYKSESLEDEEKLEIQRLEAONQERKKSQA      | 60  |
| QY | 61 GKGLTRSLAVCEESSARPGESLDOESIHJOLSFSLSOEEDKARKDSEKEXKDN         | 120 |
| Db | 61 GKGLTRSLAVCEESSARPGESLDOESIHJOLSFSLSOEEDKARKDSEKEXKDN         | 120 |
| QY | 121 KDKTSEKXIRMLSDCQOEYTDSTGIDHLEPLINTLKONSRDMLLKMEOEIIDPLAD     | 180 |
| Db | 121 KDKTSEKXIRMLSDCQOEYTDSTGIDHLEPLINTLKONSRDMLLKMEOEIIDPLAD     | 180 |
| QY | 181 NNNHKKRPQMSXYORMLVHRVAAYFGDLHNDQTKSVYINTKSTRIPEORFCHLKD      | 240 |
| Db | 181 NNNHKKRPQMSXYORMLVHRVAAYFGDLHNDQTKSVYINTKSTRIPEORFCHLKD      | 240 |
| QY | 241 EKGEESQRFILKRONSLDKEDNOSVCSQESLFEVNSRLLEBSNINCNEYKKRQLFRGN   | 300 |
| Db | 241 EKGEESQRFILKRONSLDKEDNOSVCSQESLFEVNSRLLEBSNINCNEYKKRQLFRGN   | 300 |
| QY | 301 RDGSGRTSGSRSSSENEIKMSDHQRAWGSTDSNSRNLKPMATYASFGGITVLTRGD     | 360 |
| Db | 301 RDGSGRTSGSRSSSENEIKMSDHQRAWGSTDSNSRNLKPMATYASFGGITVLTRGD     | 360 |
| QY | 361 STSSTRSGTKLSKASSESSSAGSSGSLSTRPPLDOSTPLVSGVAASGCVYPPENGG     | 420 |
| Db | 361 STSSTRSGTKLSKASSESSSAGSSGSLSTRPPLDOSTPLVSGVAASGCVYPPENGG     | 420 |
| QY | 421 GOVAPSSSTYILLPLEAATGIPPGSILLNPHGQPEVNPDDGTAIANPPTSOQPLRSAMY  | 480 |
| Db | 421 GOVAPSSSTYILLPLEAATGIPPGSILLNPHGQPEVNPDDGTAIANPPTSOQPLRSAMY  | 480 |
| QY | 481 GOSQOQPPQOQPSPOPQOQVPPQPMAGPLVTQSVQSLQASSQSVQYPAVSFPQHLLE    | 540 |
| Db | 481 GOSQOQPPQOQPSPOPQOQVPPQPMAGPLVTQSVQSLQASSQSVQYPAVSFPQHLLE    | 540 |
| QY | 541 VSPTRGHFPMRDVAIVQFGMTLSRQSSGGTTPPSPSPVYPPSLMPQPAQOQSVVIASITQ | 600 |
| Db | 541 VSPTRGHFPMRDVAIVQFGMTLSRQSSGGTTPPSPSPVYPPSLMPQPAQOQSVVIASITQ | 600 |
| QY | 601 QLPTRGFGSGGPPISQOVLQPPSPSGFQVQPPAPMAPVYVYPSGGYPTSTTQOYRPMAP  | 660 |
| Db | 601 QLPTRGFGSGGPPISQOVLQPPSPSGFQVQPPAPMAPVYVYPSGGYPTSTTQOYRPMAP  | 660 |
| QY | 661 VOYNARQSQOMFOAAQQAQAGYQVLSQOQGFQGLIGVQOQPPQSONVINNOQGTVQSVMS | 720 |
| Db | 661 VOYNARQSQOMFOAAQQAQAGYQVLSQOQGFQGLIGVQOQPPQSONVINNOQGTVQSVMS | 720 |
| QY | 721 YPTMSSYOVPMHQSGGLPQOSYQOPIMLPNAQGSLSLATGMVYCNVTPPPONNLRL     | 780 |
| Db | 721 YPTMSSYOVPMHQSGGLPQOSYQOPIMLPNAQGSLSLATGMVYCNVTPPPONNLRL     | 780 |

|    |     |                                   |     |
|----|-----|-----------------------------------|-----|
| QY | 781 | IGPHCPSSIVPVMSASCRITNCASMSNAGQVKF | 813 |
|    |     |                                   |     |
| Db | 781 | IGPHCPSSIVPVMSASCRITNCASMSNAGQVKF | 813 |

RESULT 2  
ADG17598  
ID ADG17598 standard; protein; 813 AA

AC ADG17598 ;

DT 26-FEB-2004 (first entry)

Human Br137B protein amino acid sequence.

KM neuroprotective; ophthalmological; cyostatic; cardiac; arrhythmic;  
 KM gene therapy; molecular marker; drug target; detecting; diagnosing;  
 KM staging; monitoring; prognosticking; preventing; treating;  
 KM disease predisposition; abnormal gene expression; human;  
 KM neurological disorder; visual disorder; myopathy; heart failure;  
 KM arrhythmia; cancer; Brl37B.

OS Homo sapiens.

PN WO2003085095-A2.

PD 16-OCT-2003

PF 01-APR-2003; 2003WO-US009921.

PR 11-APR-2002; 2002US-00112312.  
PR 24-MAY-2002; 2002US-003826147.  
PR 10-JUN-2002; 2002US-001647147.  
PR 13-JUN-2002; 2002US-00167631.  
PR 24-JUN-2002; 2002US-00177917.  
PR 30-JUL-2002; 2002US-0399125P.

PA (ORIG-) ORIGENE TECHNOLOGIES INC.

PI Jay G, Kovacs KF, Li X, Fan W, Shu Y, Yee A;

DR WPI; 2003-812725/76.  
DR N-PSDB; ADG17597.

PT New expressed polynucleotides and polypeptides (e.g. OTB0949) useful as  
PT molecular markers or as drug targets, for research, or for diagnosing,  
PT preventing or treating diseases associated with abnormal gene expression  
PT (e.g. cancer).

PS Claim 6; SEQ ID NO 48; 193bp; English

CC This invention relates to a novel isolated DNA sequence and the proteins  
CC encoded by them. The sequences disclosed may be useful during the  
CC development of compounds with a neuroprotective, ophthalmological,  
CC cyclostatic, cardiant or antithrombotic activity. In addition the  
CC sequences may be useful for gene therapy. Specifically claimed is an  
CC isolated DNA comprising any of the 11 fully defined sequences of 1006-  
CC 7062 bp given in the specification. The polynucleotide encodes a  
CC polypeptide comprising any of the 11 fully defined sequences of 89-1707  
CC amino acids given in the specification. The DNA and protein are useful as  
CC molecular markers, as drug targets, and for detecting, diagnosing,  
CC staging, monitoring, prognosticating, preventing, treating or determining  
CC predisposition to, various diseases and conditions associated with  
CC abnormal expression of a gene in a subject (for example neurological or  
CC visual disorders, myopathy, heart failure, arrhythmias or cancer). These  
CC may also be used in research, drug discovery, clinical medicine or  
CC forensic science. The present sequence is that of a human protein of the  
CC invention.

**SQ** Sequence 813 AA;

| Query Match | Score | DB | Length |
|-------------|-------|----|--------|
| 100.0%      | 4235  | 7  | 813    |

Best Local Similarity 100.0%; Pred. No. 3.5e-278;

1 ATGTCTGAGCAAGGAGACCTGAATCAGGCAATAGCAGAGGAAGGAGGACTGAGCAGGAG 6C

DB 438 ATGTCTGACGAGGAGGAGCTGATCAGCAATGACAGAGGAGGAGGAGCTGAGCAGAG 497  
QY 61 AGGGCCACTCCGAGAGAGCGGCACTTTTAATCAGAAAGTCTGGATGAAGAGGAAACTG 120  
DB 498 AGGGCCACTCCGAGAGAGCGGCACTTTTAATCAGAAAGTCTGGATGAAGAGGAAACTG 557  
QY 121 GAACGTCAGAGGCGGCTGAGGCTCAGATCAAGAAAGAAATCCAAAGTCAGAGCA 180  
DB 558 GAACGTCAGAGGCGGCTGAGGCTCAGATCAAGAAAGAAATCCAAAGTCAGAGCA 617  
QY 181 GGAAGAGGTTAACTGACTGCAAGCTTGTCTGTGAGGAATCTTCTGCAAGCCAGGA 240  
DB 618 GGAAGAGGTTAACTGACTGCAAGCTTGTCTGTGAGGAATCTTCTGCAAGCCAGGA 677  
QY 241 GGTGAAGTCTTCAGGATCAG 261  
DB 678 GGTGAAGTCTTCAGGATCAG 698

RESULT 2  
US-11-266-748A-147091/c  
Sequence 147091, Application US/11266748A  
Publication No. US20060134663A1  
GENERAL INFORMATION:  
APPLICANT: Harkin, Paul  
APPLICANT: Johnston, Patrick  
APPLICANT: Mulligan, Karl  
TITLE OF INVENTION: Transcriptome Microarray Technology and  
TITLE OF INVENTION: Methods of Using the Same  
FILE REFERENCE: 55815-0102 (319189)  
CURRENT APPLICATION NUMBER: US/11/266,748A  
CURRENT FILING DATE: 2005-11-03  
PRIOR APPLICATION NUMBER: EP 04105479.2  
PRIOR FILING DATE: 2004-11-03  
PRIOR APPLICATION NUMBER: EP 04105482.6  
PRIOR FILING DATE: 2004-11-03  
PRIOR APPLICATION NUMBER: EP 04105483.4  
PRIOR FILING DATE: 2004-11-03  
PRIOR APPLICATION NUMBER: EP 04105507.0  
PRIOR FILING DATE: 2004-11-03  
PRIOR APPLICATION NUMBER: EP 04105485.9  
PRIOR FILING DATE: 2004-11-03  
PRIOR APPLICATION NUMBER: EP 04105484.2  
PRIOR FILING DATE: 2004-11-03  
PRIOR APPLICATION NUMBER: US 60/662,276  
PRIOR FILING DATE: 2005-03-14  
PRIOR APPLICATION NUMBER: US 60/700,293  
PRIOR FILING DATE: 2005-07-18  
NUMBER OF SEQ ID NOS: 483996  
SOFTWARE: Patentin version 3.3  
SEQ ID NO 147091  
LENGTH: 905  
TYPE: DNA  
ORGANISM: Homo Sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (792)..(826)  
OTHER INFORMATION: n is a, c, g, or t  
US-11-266-748A-147091

Query Match 10.7%; Score 260.6; DB 8; Length 905;  
Best Local Similarity 99.6%; Pred. No. 2.2e-59;  
Matches 260; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCTGACGAGGAGGAGCTGATCAGCAATGACAGAGGAGGAGGAGCTGAGCAGAG 60  
DB 468 ATGTCTGACGAGGAGGAGCTGATCAGCAATGACAGAGGAGGAGGAGCTGAGCAGAG 409  
QY 61 AGGGCCACTCCGAGAGAGCGGCACTTTTAATCAGAAAGTCTGGATGAAGAGGAAACTG 120  
DB 408 AGGGCCACTCCGAGAGAGCGGCACTTTTAATCAGAAAGTCTGGATGAAGAGGAAACTG 349  
QY 121 GAACGTCAGAGGCGGCTGAGGCTCAGATCAAGAAAGAAATCCAAAGTCAGAGCA 180

DB 348 GAACGTCAGAGGCGGCTGAGGCTCAGATCAAGAAAGAAATCCAAAGTCAGAGCA 289  
QY 181 GGAAGAGGTTAACTGACTGCAAGCTTGTCTGTGAGGAATCTTCTGCAAGCCAGGA 240  
DB 288 GGAAGAGGTTAACTGACTGCAAGCTTGTCTGTGAGGAATCTTCTGCAAGCCAGGA 229  
QY 241 GGTGAAGTCTTCAGGATCAG 261  
DB 228 GGTGAAGTCTTCAGGATCAG 208

RESULT 3  
US-11-266-748A-381234  
Sequence 381234, Application US/11266748A  
Publication No. US20060134663A1  
GENERAL INFORMATION:  
APPLICANT: Harkin, Paul  
APPLICANT: Johnston, Patrick  
APPLICANT: Mulligan, Karl  
TITLE OF INVENTION: Transcriptome Microarray Technology and  
TITLE OF INVENTION: Methods of Using the Same  
FILE REFERENCE: 55815-0102 (319189)  
CURRENT APPLICATION NUMBER: US/11/266,748A  
CURRENT FILING DATE: 2005-11-03  
PRIOR APPLICATION NUMBER: EP 04105479.2  
PRIOR FILING DATE: 2004-11-03  
PRIOR APPLICATION NUMBER: EP 04105482.6  
PRIOR FILING DATE: 2004-11-03  
PRIOR APPLICATION NUMBER: EP 04105483.4  
PRIOR FILING DATE: 2004-11-03  
PRIOR APPLICATION NUMBER: EP 04105507.0  
PRIOR FILING DATE: 2004-11-03  
PRIOR APPLICATION NUMBER: EP 04105485.9  
PRIOR FILING DATE: 2004-11-03  
PRIOR APPLICATION NUMBER: EP 04105484.2  
PRIOR FILING DATE: 2004-11-03  
PRIOR APPLICATION NUMBER: EP 04105483.4  
PRIOR FILING DATE: 2004-11-03  
PRIOR APPLICATION NUMBER: EP 04105507.0  
PRIOR FILING DATE: 2004-11-03  
PRIOR APPLICATION NUMBER: EP 04105485.9  
PRIOR FILING DATE: 2005-03-14  
PRIOR APPLICATION NUMBER: US 60/662,276  
PRIOR FILING DATE: 2005-07-18  
NUMBER OF SEQ ID NOS: 483996  
SOFTWARE: Patentin version 3.3  
SEQ ID NO 381234  
LENGTH: 906  
TYPE: DNA  
ORGANISM: Homo Sapiens  
US-11-266-748A-381234

Query Match 10.7%; Score 260.6; DB 8; Length 906;  
Best Local Similarity 99.6%; Pred. No. 2.2e-59;  
Matches 260; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCTGACGAGGAGGAGCTGATCAGCAATGACAGAGGAGGAGGAGCTGAGCAGAG 60  
DB 263 ATGTCTGACGAGGAGGAGCTGATCAGCAATGACAGAGGAGGAGGAGCTGAGCAGAG 322  
QY 61 AGGGCCACTCCGAGAGAGCGGCACTTTTAATCAGAAAGTCTGGATGAAGAGGAAACTG 120  
DB 323 AGGGCCACTCCGAGAGAGCGGCACTTTTAATCAGAAAGTCTGGATGAAGAGGAAACTG 382  
QY 121 GAACGTCAGAGGCGGCTGAGGCTCAGATCAAGAAAGAAATCCAAAGTCAGAGCA 180  
DB 383 GAACGTCAGAGGCGGCTGAGGCTCAGATCAAGAAAGAAATCCAAAGTCAGAGCA 442  
QY 181 GGAAGAGGTTAACTGACTGCAAGCTTGTCTGTGAGGAATCTTCTGCAAGCCAGGA 240  
DB 443 GGAAGAGGTTAACTGACTGCAAGCTTGTCTGTGAGGAATCTTCTGCAAGCCAGGA 502  
QY 241 GGTGAAGTCTTCAGGATCAG 261  
DB 503 GGTGAAGTCTTCAGGATCAG 523



```
US-09-513-999C-22301
; Sequence 22301, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclet, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59, US2, REG
; CURRENT APPLICATION NUMBER: US/09/513, 999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 22301
; LENGTH: 408
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 329
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 330
; OTHER INFORMATION: d=a or g or t
US-09-513-999C-22301

Query Match
Best Local Similarity 9.2%; Score 225.6; DB 3; Length 408;
Matches 225; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 36 AGAGGAAGAGGAGTCTGAGCAGAGACGCGCATCTCCAGAGAAACGGCATTTGTAATCAGA 95
DB 1 AGAGGAAGAGGAGTCTGAGCAGAGACGCGCATCTCCAGAGAAACGGCATTTGTAATCAGA 60
QY 96 AAGCTGTGATGAAGAGGAGAACTGGAACCTGAGAGGGGGCTGAGAGCTCAAGAAATCAGA 155
DB 61 AAGCTGTGATGAAGAGGAGAACTGGAACCTGAGAGGGGGCTGAGAGCTCAAGAAATCAGA 120
QY 156 AAGAAGAAATCTCAAGTCAAGAGCAGAGAAAGTAACTGACTCGCAGCTTGTCTGTG 215
DB 121 AAGAAGAAATCTCAAGTCAAGAGCAGAGAAAGTAACTGACTCGCAGCTTGTCTGTG 180
QY 216 TGAGGAATCTTCTGCGCAGACCAAGAGGTGAAGTCTTCAGATCAG 261
DB 181 TGAGGAATCTTCTGCGCAGACCAAGAGGTGAAGTCTTCAGATCAG 226

RESULT 3
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
```

```
US-09-513-999C-22301
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpc-Fls
US-08-232-463-14

Query Match
Best Local Similarity 3.1%; Score 75.6; DB 2; Length 7218;
Matches 12; Conservative 241; Mismatches 135; Indels 0; Gaps 0;

QY 7 GAGCAAGAGACCTGGAATCAGCAATGACAGAGAGAGGAGTACGACAGAGACGCGC 66
DB 1431 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1372
QY 67 ACTCCAGAGAAACGGCATTTGTAATCAGAAAGTCTGATGAAGAGAGAACTGCACTG 126
DB 1371 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1312
QY 127 CAGAGGGGGCTGAGAGGCTCAAGATCAAGAAAGAAATCCAGTCAAGAGCAGAGAA 186
DB 1311 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1252
QY 187 GGTAACTGACTCCAGACTTGTCTGTGAGGAATCTTGTGCGCAGACGAGAGTGA 246
DB 1251 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1192
QY 247 AGCTTTCAGATCAGAAATCAATTCAATTAAGCTTTTCCAGCTTTCAGCTGCAAG 306
DB 1191 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1132
QY 307 GAGGATAAATCTAGAGAAAGATGACTGGAAGAGAAAGAAAGATTAAGAAACAAAGAT 366
DB 1131 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1072
QY 367 AAAACCTCTGAAAAACCCAGATCAGAA 394
DB 1071 RRRRRATGCAAGCTCCTGCACTGCA 1044

RESULT 4
US-09-270-767-28545/c
; Sequence 28545, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 28545
```

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2006, 09:44:00 ; Search time 11518 Seconds  
(without alignments)  
11855.792 Million cell updates/sec

Title: US-10-509-950-2

Perfect score: 2441.6  
Sequence: 1 agtccgagcagagagacc.....gtcgcagcagcraactctga 2442

Scoring table:

IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST

1: gb\_est1:\*  
2: gb\_est3:\*  
3: gb\_est4:\*  
4: gb\_est5:\*  
5: gb\_est6:\*  
6: gb\_est7:\*  
7: gb\_est8:\*  
8: gb\_est9:\*  
9: gb\_est10:\*  
10: gb\_est11:\*  
11: gb\_est12:\*  
12: gb\_est13:\*  
13: gb\_est14:\*  
14: gb\_est15:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1          | 1944   | 79.6        | 2505   | 6     | HSM801390   |
| 2          | 1815.8 | 74.4        | 2914   | 6     | AK034376    |
| 3          | 1724.8 | 70.6        | 3180   | 6     | AK032417    |
| 4          | 1641.2 | 67.2        | 2991   | 6     | AK032543    |
| 5          | 938    | 38.4        | 2806   | 6     | AK137974    |
| 6          | 910.4  | 37.3        | 1550   | 8     | CO635664    |
| 7          | 757.6  | 31.0        | 845    | 2     | BI553063    |
| 8          | 748    | 30.6        | 925    | 10    | DV909010    |
| 9          | 743    | 30.4        | 1053   | 3     | BM926937    |
| 10         | 742.6  | 30.4        | 929    | 10    | DV909014    |
| 11         | 695    | 28.5        | 947    | 10    | DV879970    |
| 12         | 662.4  | 27.3        | 743    | 9     | DN991171    |
| 13         | 641    | 26.3        | 991    | 4     | BY761855    |
| 14         | 641    | 26.3        | 991    | 4     | BY761855    |
| 15         | 630    | 25.8        | 1864   | 6     | AK029467    |
| 16         | 620.2  | 25.4        | 922    | 6     | AK153892    |
| 17         | 615    | 25.2        | 749    | 4     | BU164142    |
| 18         | 605.8  | 24.8        | 609    | 9     | DA529811    |
| 19         | 600.4  | 24.6        | 1056   | 10    | DV797399    |

|    |       |      |      |   |          |
|----|-------|------|------|---|----------|
| 20 | 590.8 | 24.2 | 866  | 4 | EX461002 |
| 21 | 588.8 | 24.1 | 753  | 5 | CJ464930 |
| 22 | 586   | 24.0 | 782  | 4 | EX431175 |
| 23 | 571.4 | 23.4 | 574  | 9 | DA487336 |
| 24 | 571   | 23.4 | 571  | 9 | DA494904 |
| 25 | 569.6 | 23.3 | 740  | 5 | CJ439713 |
| 26 | 565.6 | 23.2 | 837  | 3 | BQ217720 |
| 27 | 562.8 | 23.1 | 926  | 3 | BP160144 |
| 28 | 552.8 | 22.6 | 634  | 9 | DN427693 |
| 29 | 551.2 | 22.6 | 629  | 5 | CK003565 |
| 30 | 549.4 | 22.5 | 563  | 9 | DA764999 |
| 31 | 547.8 | 22.4 | 575  | 9 | DA179590 |
| 32 | 540.2 | 22.1 | 654  | 7 | BB626408 |
| 33 | 540.2 | 22.1 | 891  | 3 | BP160316 |
| 34 | 538.2 | 22.0 | 1946 | 6 | AK045507 |
| 35 | 536.6 | 22.0 | 711  | 3 | BP162992 |
| 36 | 534.6 | 21.9 | 1688 | 6 | AK050613 |
| 37 | 534.4 | 21.9 | 765  | 3 | BO179621 |
| 38 | 526.4 | 21.6 | 728  | 3 | BM951156 |
| 39 | 516   | 21.1 | 566  | 9 | DA260904 |
| 40 | 516   | 21.1 | 595  | 9 | DB172765 |
| 41 | 511.8 | 21.0 | 806  | 5 | CK460159 |
| 42 | 506.4 | 20.7 | 728  | 3 | BU342596 |
| 43 | 506.4 | 20.7 | 800  | 3 | BU604247 |
| 44 | 493   | 20.2 | 575  | 9 | DA389805 |
| 45 | 491.2 | 20.1 | 640  | 7 | BB622926 |

#### ALIGNMENTS

RESULT 1  
HSM801390  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

Homo sapiens (human)  
Homo sapiens  
Homo sapiens

#### REFERENCE

#### AUTHORS

#### CONSTRM

#### TITLE

#### JOURNAL

#### COMMENT

#### FEATURES

source  
1. 2505  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="RZPD:DKFZp566N1047"  
/db\_xref="taxon:9606"  
/clone="DKFZp566N1047"  
/issue\_type="kidney"  
/clone\_lib="566 (synonym: hfk2). Vector pAMP1, host  
Xl-2Blue; sites NciI + SalI"  
/dev\_stage="fetal"  
/note="cyclic AMP-regulated phosphoprotein, 21 kD,



gene N-terminus truncated, differentially spliced"  
1..2505  
/gene="DKFZp566N1047"  
cms <3..1958  
/gene="DKFZp566N1047"  
/product="hypothetical protein"  
/protein\_id="CAB61414.1"  
/db\_xref="GI:6453596"  
/db\_xref="GOA:O9UF93"  
/translation="RDRMLLKMOEIIIDTADNNHYKKPEPOMSSYORMLVHVAAY  
FGLDHVDQKGVIIINKSTSTRIPRORPCHEIKDEKESORRFLIKRNSIIDKED  
NOSVCSOESLFEVNRILEDNINCHETKTKKQLRGNRDSGRSGRSQSSSENLKRS  
DHQWMSSTDSNRLKPKMTKTSFGIIVLRDSTSRSTKSSSENLKRS  
SAGSGLSRTHPLQSTPLVSGVAAGSPCCVPYENGIGQVAPSTSYILPLEA  
TGIPLGSIILNPHGTQFVNPDPGTPALYNPTSQPLRSAMQSQOQPPQOQPSIP  
QOQVOPQPMAGPLVTSVOGLQASSQVOYPAVSFPPLHPSPTOHFMRDVA  
TORGMKLSRSGGTEPSPGVPYSSLMOPAPQOSPVIASITGQQLPTGFSGSP  
PIQOYIOPPPSPQFYOQPPAPMPYRYSQYPISTTQYRPMAPVYNQSRQ  
MPOAQOAGTQPVLSGQGFQGLGVQOQPOQSVINNOGTFVQSWVSPTMSSTO  
VPMTOGQGLPQOQYQPIMLPNOGGSPLATGMVPCNVTPPTPQNNRLIGPHCP  
STVPMVASCRNCA SMSNAGWQVKE"

## ORIGIN

Query Match 79.6%; Score 1944; DB 6; Length 2505;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1957; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

483 CAGGACAGAGTACTTTTGAATAATGAGCAGGAATATTGATTTCAATTCGTCGACAA 542.  
DB 2 CAGGACAGAGTACTTTTGAATAATGAGCAGGAATATTGATTTCAATTCGTCGACAA 61  
QY 543 CAATATCATTTATTAATAAGTTCCCTCAGATGTCATCGTATCAGAGATGCTTGTCCATCG 602  
DB 62 CAATATCATTTATTAATAAGTTCCCTCAGATGTCATCGTATCAGAGATGCTTGTCCATCG 121  
QY 603 AGTGGCAGCTTATTTTGGATTGATTCACATGTGTGATCAACAGAGAAATCTGTATCAT 662  
DB 122 AGTGGCAGCTTATTTTGGATTGATTCACATGTGTGATCAACAGAGAAATCTGTATCAT 181  
QY 663 CAAACAAGACGACGACACAGAAATACAGAGCAAAAGTTTGTGATCAATTTAAAGATGA 722  
DB 182 CAAACAAGACGACGACACAGAAATACAGAGCAAAAGTTTGTGATCAATTTAAAGATGA 241  
QY 723 AAAAGTGAAGATCCAGAAACCGGTTTATCTTGAAACGAGATTACTAGATATTGATTA 782  
DB 242 AAAAGTGAAGATCCAGAAACCGGTTTATCTTGAAACGAGATTACTAGATATTGATTA 301  
QY 783 AGAAGACATCATGTCATGTTGCTCCAGAAAGCTTTTGTGAAACAGTAGGCTCTT 842  
DB 302 AGAAGACATCATGTCATGTTGCTCCAGAAAGCTTTTGTGAAACAGTAGGCTCTT 358  
QY 843 GGAAGACATTAATGATGAGACCTTAATAAGAAAGACAGCTCTTGGGGGCAACAG 902  
DB 359 GGAAGACATTAATGATGAGACCTTAATAAGAAAGACAGCTCTTGGGGGCAACAG 418  
QY 903 AGATGCTCAGAGGAAACATCTGGAGTCCACAGAGACGCTCAGAAAATGATCAAGTG 962  
DB 419 AGATGCTCAGAGGAAACATCTGGAGTCCACAGAGACGCTCAGAAAATGATCAAGTG 478  
QY 963 GTCTGACCAACAAAGGGCTTGAGACAGACAGATCTCCGACAGTTCCAAACCGCAATCTTAAA 1022  
DB 479 GTCTGACCAACAAAGGGCTTGAGACAGACAGATCTCCGACAGTTCCAAACCGCAATCTTAAA 538  
QY 1023 GCCCGCATGACCAAGAGCGGAGTTTGGGGGATCATCAGGTCGTGACAGAGGGGTGACAG 1082  
DB 539 GCCCGCATGACCAAGAGCGGAGTTTGGGGGATCATCAGGTCGTGACAGAGGGGTGACAG 598  
QY 1083 CACTTCCAGTACTAGAGTACCGGAAAGCTGTCCAAAGCAGGTTCCAGTCTTCCAGACAG 1142  
DB 599 CACTTCCAGTACTAGAGTACCGGAAAGCTGTCCAAAGCAGGTTCCAGTCTTCCAGACAG 658

QY 1143 TGCAGGCTCCTCAGATCGCTGTCCGCAACCCATCCACTCTCCAGAGACACACCCCTAGT 1202  
DB 659 TGCAGGCTCCTCAGATCGCTGTCCGCAACCCATCCACTCTCCAGAGACACACCCCTAGT 718  
QY 1203 CTCAGGTTGTGACGTGCTCTCCAGAGCTGTGCTTATCCAGAGATGAAATAGGAGG 1262  
DB 719 CTCAGGTTGTGACGTGCTCTCCAGAGCTGTGCTTATCCAGAGATGAAATAGGAGG 778  
QY 1263 CCAGGTTGCTCCCAAGACACAGCTCATCTCTCTTCCACTTGAAGCTGCAACAGCAT 1322  
DB 779 CCAGGTTGCTCCCAAGACACAGCTCATCTCTCTTCCACTTGAAGCTGCAACAGCAT 838  
QY 1323 CCGGCGGAAAGCATCCTTTATATCAACACAGGCGAGCCCTTGTGAATCCGATG 1382  
DB 839 CCGGCGGAAAGCATCCTTTATATCAACACAGGCGAGCCCTTGTGAATCCGATG 898  
QY 1383 AACTCTGCAATATTAACACCAACCAACAGTACAGAGCCCTGGAAGCGCATGTGG 1442  
DB 899 AACTCTGCAATATTAACACCAACCAACAGTACAGAGCCCTGGAAGCGCATGTGG 958  
QY 1443 GCAGTCCCAACAGACCGGCTCAACAGACAGCCCTCCCGACGCCCAACAGAGTCCA 1502  
DB 959 GCAGTCCCAACAGACCGGCTCAACAGACAGCCCTCCCGACGCCCAACAGAGTCCA 1018  
QY 1503 GCCACGCGACGACAGATGAGGAGGCTCTGTGATCACTAGTCTGTCCAGGGGCTGACAGG 1562  
DB 1019 GCCACGCGACGACAGATGAGGAGGCTCTGTGATCACTAGTCTGTCCAGGGGCTGACAGG 1078  
QY 1563 TTCCTCCAGTACAGTCAATATCCGAGAGTCTCTTTTCTCCCGACAGCTCTTACTGT 1622  
DB 1079 TTCCTCCAGTACAGTCAATATCCGAGAGTCTCTTTTCTCCCGACAGCTCTTACTGT 1138  
QY 1623 GTCTCCACGACAGACTTTTCCATGAGATGATGTGGCAACAAGTTTGGCCAGATGAC 1682  
DB 1139 GTCTCCACGACAGACTTTTCCATGAGATGATGTGGCAACAAGTTTGGCCAGATGAC 1198  
QY 1683 CTTGAGCGGAGTCTCTGGGGGAGACTCTGTAACCCCATCAGGTCCTGTCTACCATC 1742  
DB 1199 CTTGAGCGGAGTCTCTGGGGGAGACTCTGTAACCCCATCAGGTCCTGTCTACCATC 1258  
QY 1743 CTCCTTATGCAACAGCGGCGCCAGACAGCCCACTATGATGTGCTTCAAGGCGACGA 1802  
DB 1259 CTCCTTATGCAACAGCGGCGCCAGACAGCCCACTATGATGTGCTTCAAGGCGACGA 1318  
QY 1803 GCTTCTCAAGAGGATCTCAGGCTGTGGCTCTCCATCTCCAGAGAGTCTTCCAGCC 1862  
DB 1319 GCTTCTCAAGAGGATCTCAGGCTGTGGCTCTCCATCTCCAGAGAGTCTTCCAGCC 1378  
QY 1863 CCGTCCCTCAACAGAGGATCTGAGCAACAGGCTCCGCTGACAGATGCTGTATATTA 1922  
DB 1379 CCGTCCCTCAACAGAGGATCTGAGCAACAGGCTCCGCTGACAGATGCTGTATATTA 1438  
QY 1923 TTACCATCTGTCAATACCTTACCTTCAACACAGCAACAGTACCGGCTTATGCGCCGGT 1982  
DB 1439 TTACCATCTGTCAATACCTTACCTTCAACACAGCAACAGTACCGGCTTATGCGCCGGT 1458  
QY 1983 TCACTACACGCTCAGAGAGTCAACAGATGCCACAGGCAACAGAAAGAGTTTACA 2042  
DB 1499 TCACTACACGCTCAGAGAGTCAACAGATGCCACAGGCAACAGAAAGAGTTTACA 1558  
QY 2043 GCCAGCTTGTCTGTCAACAGGATTTCAAGGCTTATAGAGTGTGAGCAGGACGACTCA 2102  
DB 1559 GCCAGCTTGTCTGTCAACAGGATTTCAAGGCTTATAGAGTGTGAGCAGGACGACTCA 1618  
QY 2103 GAGTCAAGACGTATTAATTAACAACAAGAACTCCGGTCAAAAGCTGATGTTTCTTA 2162  
DB 1619 GAGTCAAGACGTATTAATTAACAACAAGAACTCCGGTCAAAAGCTGATGTTTCTTA 1678  
QY 2163 CCCAACAATGCTCTTATAGAGTCCCAATGACCAAGGCTTCTCAAGGATGCCCCAGCA 2222  
DB 1679 CCCAACAATGCTCTTATAGAGTCCCAATGACCAAGGCTTCTCAAGGATGCCCCAGCA 1738  
QY 2223 GTCAATCAACACAGCAATCATGCTTAAACAGGACAGTCAAGGCTCACTCCAGGCAC 2282



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2006, 09:44:00 ; Search time 1406 Seconds

(without alignments)  
12109.695 Million cell updates/sec

Title: US-10-509-950-2

Perfect score: 2441.6  
Sequence: 1 atgtctgagcaagagagacct.....gttcgacggtcaattctga 2442

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database: N\_Geneseq\_B:\*

- 1: geneseqn1980s:\*
- 2: geneseqn1990s:\*
- 3: geneseqn2000s:\*
- 4: geneseqn2001as:\*
- 5: geneseqn2001bs:\*
- 6: geneseqn2002as:\*
- 7: geneseqn2002bs:\*
- 8: geneseqn2003as:\*
- 9: geneseqn2003bs:\*
- 10: geneseqn2003cs:\*
- 11: geneseqn2003ds:\*
- 12: geneseqn2004as:\*
- 13: geneseqn2004bs:\*
- 14: geneseqn2005s:\*
- 15: geneseqn2006s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | ID | Description          |
|------------|--------|-------------|--------|----|----------------------|
| 1          | 2441.6 | 100.0       | 3212   | 10 | ADD89837 Human CAM   |
| 2          | 2441.2 | 100.0       | 3332   | 10 | ADG17597 Human Brl   |
| 3          | 2324.8 | 95.2        | 3369   | 10 | ADG17593 Human TAR   |
| 4          | 2311.2 | 94.7        | 3272   | 10 | ADG17599 Human Brl   |
| 5          | 2287.2 | 93.7        | 3374   | 10 | ADG17595 Human Brl   |
| 6          | 2200.8 | 90.1        | 2813   | 6  | ABQ60920 Human Brl   |
| 7          | 1746.4 | 71.5        | 2548   | 5  | AA566277 DKFZ566n    |
| 8          | 1366.2 | 56.0        | 1942   | 5  | AA566275 DNA encod   |
| 9          | 750.2  | 30.7        | 1305   | 5  | AA566275 DNA encod   |
| 10         | 650.4  | 26.6        | 1488   | 5  | AA566275 DNA encod   |
| 11         | 650.4  | 26.6        | 1488   | 5  | AA566275 DNA encod   |
| 12         | 413.2  | 14.8        | 468    | 5  | ABAI3329 Human encod |
| 13         | 362.4  | 14.8        | 4340   | 13 | ABAI3329 Human encod |
| 14         | 353.4  | 14.5        | 625    | 5  | ABAI2203 Human nar   |
| 15         | 308.8  | 12.6        | 4330   | 8  | ACCT2057 Human nar   |
| 16         | 308.8  | 12.6        | 4331   | 10 | ACCT2057 Human nar   |
| 17         | 274.8  | 11.3        | 3839   | 8  | ACCS9933 Human IRA   |
| 18         | 274.8  | 11.3        | 4373   | 10 | ADG37200 Nuclear f   |

|    |       |      |      |    |                    |
|----|-------|------|------|----|--------------------|
| 19 | 262.2 | 10.7 | 2014 | 10 | ADB63612 Human CDN |
| 20 | 260.6 | 10.7 | 352  | 6  | ABA04824 Human hAR |
| 21 | 260.6 | 10.7 | 1006 | 10 | ADG17601 Human Brl |
| 22 | 260.6 | 10.7 | 1048 | 10 | ADD29568 Human tum |
| 23 | 260.6 | 10.7 | 4276 | 8  | ACCT2056 BCUI147A  |
| 24 | 259   | 10.6 | 876  | 3  | AAA47416 Sequence  |
| 25 | 257.4 | 10.5 | 371  | 2  | AAQ60619 Human bra |
| 26 | 257.4 | 10.5 | 371  | 5  | AA566273 DNA encod |
| 27 | 257.4 | 10.5 | 371  | 5  | AA566273 DNA encod |
| 28 | 255.8 | 10.5 | 4272 | 15 | AEF74808 Human pol |
| 29 | 233.8 | 9.6  | 579  | 12 | ACH71408 Human gen |
| 30 | 231.2 | 9.5  | 274  | 12 | ACH85164 Human gen |
| 31 | 229.4 | 9.4  | 567  | 5  | AA566274 DNA encod |
| 32 | 225.6 | 9.2  | 408  | 3  | AAC18226 Human sec |
| 33 | 207.2 | 8.5  | 308  | 5  | AAQ61132 Human bra |
| 34 | 207.2 | 8.5  | 308  | 5  | AA569171 DNA encod |
| 35 | 207.2 | 8.5  | 308  | 5  | AA566272 DNA encod |
| 36 | 206.6 | 8.5  | 646  | 13 | ADQ52802 Novel can |
| 37 | 203.2 | 8.3  | 1305 | 5  | AA566275 DNA encod |
| 38 | 203.2 | 8.3  | 1942 | 5  | AA569174 DNA encod |
| 39 | 177   | 7.2  | 404  | 9  | ACH46873 Human inf |
| 40 | 160   | 6.6  | 1030 | 10 | ADB53792 Primary r |
| 41 | 160   | 6.6  | 1030 | 10 | ADB53792 Primary r |
| 42 | 150.8 | 6.2  | 285  | 6  | ABN77809 Human ORF |
| 43 | 136.2 | 5.6  | 2906 | 10 | ACH71296 Nuclear f |
| 44 | 134.8 | 5.5  | 510  | 12 | ACH71623 Human gen |
| 45 | 129   | 5.3  | 302  | 12 | ACH85379 Human gen |

## ALIGNMENTS

|          |  |
|----------|--|
| RESULT 1 | ADD89837 standard; cDNA; 3212 BP.  |
| ID       | ADD89837   |
| AC       | ADD89837   |
| XX       | 29-JAN-2004 (first entry)  |
| XX       | Human cAMP-regulated phosphoprotein hTARPP cDNA.                         |
| XX       | Human, hTARPP, TARPP, neurodegenerative disease; Alzheimer's disease;    |
| KW       | gene therapy; neuroprotective; nootropic; transgenic; gene; ss.          |
| OS       | Homo sapiens.  |
| XX       | Key  |
| PH       | Location/Qualifiers  |
| FT       | 240..2681  |
| FT       | /*tag= a   |
| FT       | /product= "hTARPP"   |
| FT       | /note= "The CDS is also claimed in Claim 3"                              |
| XX       | MO2003083482-A1.   |
| XX       | 09-OCT-2003.   |
| XX       | 01-APR-2003; 2003MO-EP003364.  |
| XX       | 02-APR-2002; 2002EP-0007522.   |
| XX       | 02-APR-2002; 2002US-0368970P.  |
| XX       | (EVOT-) EVOTEC NEUROSCIENCES GMBH.                                       |
| XX       | Hipfler R, Hanes J, Von Der Kammer H, Pohlner J;                         |
| XX       | WPI, 2003-804100/75.   |
| XX       | P-PSDB; ADD89836.  |
| XX       | New nucleic acid useful for preparing a composition for treating,        |
| XX       | preventing, diagnosing, prognosticating or monitoring the progression of |
| XX       | a neurodegenerative disease, e.g. Alzheimer's disease.                   |

PS Claim 3; SEQ ID NO 3; 69pp; English.

The present sequence, the coding region of which is also claimed, is that of cDNA encoding human cAMP-regulated phosphoprotein hTRAPP (or hTRAPP), a previously undescribed human isoform of murine cAMP regulated phosphoprotein ARPP-21. The cDNA comprises an alignment of human expressed sequence tag clones that were obtained from a database screening using a differentially expressed human cDNA fragment AB089838. The invention discloses the differential expression of hTRAPP in specific brain regions of Alzheimer's disease patients. On the basis of this finding, a method is provided for diagnosing or prognosticating a neurodegenerative disease, in particular Alzheimer's disease, or for determining whether a subject is at increased risk of developing such a disease. The gene encoding hTRAPP is used in therapeutic and prophylactic methods for treating or preventing Alzheimer's disease and related neurodegenerative disorders. A method of screening for modulating agents of neurodegenerative diseases is also provided. Transgenic animals comprising a non-native gene sequence coding for hTRAPP are useful for screening, testing or validating compounds, agents or modulators, and developing diagnostics or therapeutics for treating neurodegenerative disease, particularly Alzheimer's disease. The encoded hTRAPP protein is useful as a diagnostic target for detecting, or screening target for reagents or compounds for preventing, treating or ameliorating a neurodegenerative disease, preferably Alzheimer's disease.

SQ Sequence 3212 BP; 938 A; 795 C; 745 G; 733 T; 0 U; 1 Other;

|                       |        |              |       |             |
|-----------------------|--------|--------------|-------|-------------|
| Query Match           | 100.0% | Score 2441.6 | DB 10 | Length 3222 |
| Best Local Similarity | 100.0% | Pred. No. 0  |       |             |
| Matches 2442          | 0      | Mismatches   | 0     | Gaps 0      |

|    |     |   |     |
|----|-----|---|-----|
| OY | 1   | ATGCTGAGCAGAGGAGACCTGAAATCAGGCAATTCAGAGGAAGGAGGACCTGAGCCAGAG  | 60  |
| Db | 240 | ATGCTGAGCAGAGGAGACCTGAAATCAGGCAATTCAGAGGAAGGAGGACCTGAGCCAGAG  | 239 |
| OY | 61  | ACGGCACTCCAGAGAACGGCACTTGTTAAATCAGAAAGTCTGGATGAGGAGAACTG      | 120 |
| Db | 300 | ACGGCCACTCCAGAGAACGGCACTTGTTAAATCAGAAAGTCTGGATGAGGAGAACTG     | 359 |
| OY | 121 | GAACTGCAAGAGCGCTGAGAGGCTCAAGATTCAGAAAGAAAGAAATTCAGATCAGGAGCA  | 180 |
| Db | 360 | GAACTGCAAGAGCGCTGAGAGGCTCAAGATTCAGAAAGAAAGAAATTCAGATCAGGAGCA  | 419 |
| OY | 181 | GGAAGAAAGGTAACTGACTCGCAGCTCTGCTGTGTGAGGAATCTTCTGCAGACCAAGA    | 240 |
| Db | 420 | GGAAGAAAGGTAACTGACTCGCAGCTCTGCTGTGTGTGAGGAATCTTCTGCAGACCAAGA  | 479 |
| OY | 241 | GGTGAAAGCTTCAAGATCAGAAATCAATTCATTCACGTTTCCAGTTCACAGCTG        | 300 |
| Db | 480 | GGTGAAGCTTCAAGATCAGAAATCAATTCATTCACGTTTCCAGTTCACAGCTG         | 539 |
| OY | 301 | CAGAGAGAGATAAATCTAGAAAGATGACTCTGAAAGAGAAAAAGAAAGATTAATAAC     | 360 |
| Db | 540 | CAGAGAGAGATAAATCTAGAAAGATGACTCTGAAAGAGAAAAAGAAAGATTAATAAC     | 599 |
| OY | 361 | AAAGATAAACTCTGAAAAACCCAGATGAGAATGTTATCAAAAGTTGCAGCCAGAA       | 420 |
| Db | 600 | AAAGATAAACTCTGAAAAACCCAGATGAGAATGTTATCAAAAGTTGCAGCCAGAA       | 659 |
| OY | 421 | TACACGAGATTCTACAGGCAATAGACTTAACAGATTTCTGATTAACATTAAGAAATAT    | 480 |
| Db | 660 | TACACGAGATTCTACAGGCAATAGACTTAACAGATTTCTGATTAACATTAAGAAATAT    | 719 |
| OY | 481 | TCCAGGAGCAGAGATGATACCTTTGAAAAATGAGCAGAGAAATTAATGATTTCAATTCGAC | 540 |
| Db | 720 | TCCAGGAGCAGAGATGATACCTTTGAAAAATGAGCAGAGAAATTAATGATTTCAATTCGAC | 779 |
| OY | 541 | AACAAATAATCATATAAAAAAGTTCCCTCAGANTGTCAATCGATACAGAGATGCTTGCCAT | 600 |
| Db | 780 | AACAAATAATCATATAAAAAAGTTCCCTCAGANTGTCAATCGATACAGAGATGCTTGCCAT | 839 |
| OY | 601 | CGAGTGGCAGCTTATTTGGATTGGATCACAAATGGATCAACAGAGAAATCTGTATC      | 660 |

3

|    |      |  |      |
|----|------|--|------|
| QY | 181  | GGAAAAGTAACTGACCTCCGAGVCTTCTGTCTGTGAGAAATCTTCTCCAGACAGGA         | 240  |
| Db | 181  | GGAAAAGGTAACTGACCTCGAGCTTCTGTCTGTGAGAAATCTTCTCCAGACAGGA          | 240  |
| QY | 241  | GGTGAAAGTCTTCAGGATCAGAAATCAATTCATTAAACGCTTCAGTTTCCAGCCTG         | 300  |
| Db | 241  | GGTGAAAGTCTTCAGGATCAGAAATCAATTCATTAAACGCTTCAGTTTCCAGCCTG         | 300  |
| QY | 301  | CAAGAGAGATAAATCTAGGAAAGATGACTCTGAAAGAGAAAAAGAAAGATPAAAAC         | 360  |
| Db | 301  | CAAGAGAGATAAATCTAGGAAAGATGACTCTGAAAGAGAAAAAGAAAGATPAAAAC         | 360  |
| QY | 361  | AAAGTAAATACCTCTGAAAAACCCAAAGTCAGATGTTATATCAAAAAGATTCGACCCAGA     | 420  |
| Db | 361  | AAAGTAAATACCTCTGAAAAACCCAAAGTCAGATGTTATATCAAAAAGATTCGACCCAGA     | 420  |
| QY | 421  | TACACGGATTCTACAGGATAGACTTAACGAGTTCTGATTAAACATTTAAAGATPAT         | 480  |
| Db | 421  | TACACGGATTCTACAGGATAGACTTAACGAGTTCTGATTAAACATTTAAAGATPAT         | 480  |
| QY | 481  | TCGAGGACAGAGATGATACCTTTTGAAAAATGAGCAGGAAATTAATGATTTCTGCTGAC      | 540  |
| Db | 481  | TCGAGGACAGAGATGATACCTTTTGAAAAATGAGCAGGAAATTAATGATTTCTGCTGAC      | 540  |
| QY | 541  | AACATTAATCATTTATAAAAAGTTCCCTCAGATGTCATGTTATCAGAGATCCTGTCCAT      | 600  |
| Db | 541  | AACATTAATCATTTATAAAAAGTTCCCTCAGATGTCATGTTATCAGAGATCCTGTCCAT      | 600  |
| QY | 601  | CGAGTGGAGCTTATTTTGGATTCGATGCAATGTTGGAATCAACAGGAAATCTGTATTC       | 660  |
| Db | 601  | CGAGTGGAGCTTATTTTGGATTCGATGCAATGTTGGAATCAACAGGAAATCTGTATTC       | 660  |
| QY | 661  | ATCAACAAGACAGCAGCAGCAGATACCAAGCAAAGTTTGTGAACATTTAAAAAGAT         | 720  |
| Db | 661  | ATCAACAAGACAGCAGCAGCAGATACCAAGCAAAGTTTGTGAACATTTAAAAAGAT         | 720  |
| QY | 721  | GAAAAAGGTGAAGAAATCCCAAGAGCGGTTTATCTTGAAGCAGATTACTTAGATTTGAT      | 780  |
| Db | 721  | GAAAAAGGTGAAGAAATCCCAAGAGCGGTTTATCTTGAAGCAGATTACTTAGATTTGAT      | 780  |
| QY | 781  | AAAGAAGCAATCACTGCTTGTGCTCCAGGAAAGCCTTTTGTGGAATAACATAGAGCTC       | 840  |
| Db | 781  | AAAGAAGCAATCACTGCTTGTGCTCCAGGAAAGCCTTTTGTGGAATAACATAGAGCTC       | 840  |
| QY | 841  | TTGGNAGACATTAACATATGCAATGAGACTATTAAGAAACAACAGCTCTTCGGGGCAAC      | 900  |
| Db | 841  | TTGGNAGACATTAACATATGCAATGAGACTATTAAGAAACAACAGCTCTTCGGGGCAAC      | 900  |
| QY | 901  | AGAGATGCTCAGGAGAAACATCTTGGAGTCCGACAGACAGCTCAGAAATGAATCTCAG       | 960  |
| Db | 901  | AGAGATGCTCAGGAGAAACATCTTGGAGTCCGACAGACAGCTCAGAAATGAATCTCAG       | 960  |
| QY | 961  | TGCTGTGACCAACAAAGGGCCTTGAGACGACACAGCTCCGACAGTTCCAACCGCACTTA      | 1020 |
| Db | 961  | TGCTGTGACCAACAAAGGGCCTTGAGACGACACAGCTCCGACAGTTCCAACCGCACTTA      | 1020 |
| QY | 1021 | AAAGCCCCGCAATGACCAAGACGGCGAGTTTGTGGGGCATCAAGGTGCTGACAGGGGGTAC    | 1080 |
| Db | 1021 | AAAGCCCCGCAATGACCAAGACGGCGAGTTTGTGGGGCATCAAGGTGCTGACAGGGGGTAC    | 1080 |
| QY | 1081 | AGCATTCACAGTACTTAGAGATACCGGGAAGCTGTCTCAAAAGCAGGTTCCGACTCTTCCAGC  | 1140 |
| Db | 1081 | AGCATTCACAGTACTTAGAGATACCGGGAAGCTGTCTCAAAAGCAGGTTCCGACTCTTCCAGC  | 1140 |
| QY | 1141 | AGTGCAGGCTCTCAGAGATGCGCTGTCCCGACCCATTCACCTTCCAAAGACACCCCTA       | 1200 |
| Db | 1141 | AGTGCAGGCTCTCAGAGATGCGCTGTCCCGACCCATTCACCTTCCAAAGACACCCCTA       | 1200 |
| QY | 1201 | GTTCTCAGGTGTGGCAGCTGGGCTCTTCCAGGCTGTGTGCTTTATCCAGAGAAATGGAATAGGG | 1260 |
| Db | 1201 | GTTCTCAGGTGTGGCAGCTGGGCTCTTCCAGGCTGTGTGCTTTATCCAGAGAAATGGAATAGGG | 1260 |
| QY | 1261 | GGCCAGGTGTCTCCAGCAGCAGCAGCTACATCCTCTTCCACTTGAAGCTGCAACAGCC       | 1320 |

|    |      |   |      |
|----|------|---|------|
| Db | 1261 | GGCCAGGTTGCTCCACAGACACAGCTTACCTCTTCCACTTGAAGTCGACCAAGCC       | 1320 |
| QY | 1321 | ATCCCGCGTGAAGACATCTTCTTTAATCAACAACAGGCGAGCCCTTGTGAATCCCGAT    | 1380 |
| Db | 1321 | ATCCCGCGTGAAGACATCTTCTTTAATCAACAACAGGCGAGCCCTTGTGAATCCCGAT    | 1380 |
| QY | 1381 | GGAACTCTCTGCAATATATCAACCCACCCACAGTCAAGACCCCTCTGGCAAGCCCATGCTG | 1440 |
| Db | 1381 | GGAACTCTCTGCAATATATCAACCCACCCACAGTCAAGACCCCTCTGGCAAGCCCATGCTG | 1440 |
| QY | 1441 | GGGCAAGTCCCAACAGCAGCGGCCAACAGACAGAGCCCTCCCGAGCCGCCAACAGCAGTCC | 1500 |
| Db | 1441 | GGGCAAGTCCCAACAGCAGCGGCCAACAGACAGAGCCCTCCCGAGCCGCCAACAGCAGTCC | 1500 |
| QY | 1501 | CAGCCACCGCAGACAGATGAGAGGCCCTCTGTGTCATCAGTCTGTCAAGGGAGCTGACAG  | 1560 |
| Db | 1501 | CAGCCACCGCAGACAGATGAGAGGCCCTCTGTGTCATCAGTCTGTCAAGGGAGCTGACAG  | 1560 |
| QY | 1561 | GCTTCTCTCCAGTCAATATCCGGCAGTCTCTTCTCTCCCGACACCTCTCTACT         | 1620 |
| Db | 1561 | GCTTCTCTCCAGTCAATATCCGGCAGTCTCTTCTCTCCCGACACCTCTCTACT         | 1620 |
| QY | 1621 | GTTGCTCCACAGCAGCCTTCCATGAGATGATGTGCAACAGTTTGGCCAAATG          | 1680 |
| Db | 1621 | GTTGCTCCACAGCAGCCTTCCATGAGATGATGTGCAACAGTTTGGCCAAATG          | 1680 |
| QY | 1681 | ACCCTGAGCGGCGAGTCTCTCGGGGAGACTCTTGAAACCCCATCAGGTCTCTGTACCA    | 1740 |
| Db | 1681 | ACCCTGAGCGGCGAGTCTCTCGGGGAGACTCTGTGAAACCCCATCAGGTCTCTGTACCA   | 1740 |
| QY | 1741 | TCCTCTCCCTTATGCCACAGCCGGGCCAGACGCCAGCTATGTATCGCTCTACAGGCCAG   | 1800 |
| Db | 1741 | TCCTCTCCCTTATGCCACAGCCGGGCCAGACGCCAGCTATGTATCGCTCTACAGGCCAG   | 1800 |
| QY | 1801 | CAGCTTCCCTTACAGAGGATTTCTCAGGCTCTGGCCCTCCACTTCCCGACAGGTCCTCAG  | 1860 |
| Db | 1801 | CAGCTTCCCTTACAGAGGATTTCTCAGGCTCTGGCCCTCCACTTCCCGACAGGTCCTCAG  | 1860 |
| QY | 1861 | CCCCCTCCCTCACACAGAGGATTTCTGTGCAACAGCTCTCGGCTGACAGATGCTGTATAT  | 1920 |
| Db | 1861 | CCCCCTCCCTCACACAGAGGATTTCTGTGCAACAGCTCTCGGCTGACAGATGCTGTATAT  | 1920 |
| QY | 1921 | TATTTACCCATCTGTGTCAGTACCTTACTCTCAACGCAAGTACGGGCCCATGAGCCCG    | 1980 |
| Db | 1921 | TATTTACCCATCTGTGTCAGTACCTTACTCTCAACGCAAGTACGGGCCCATGAGCCCG    | 1980 |
| QY | 1981 | GTTTCAGTACAAAGCTCAGAGAGTCAACAATGTCACAGGACAGCAGCAAGCAGGTTAC    | 2040 |
| Db | 1981 | GTTTCAGTACAAAGCTCAGAGAGTCAACAATGTCACAGGACAGCAGCAAGCAGGTTAC    | 2040 |
| QY | 2041 | CAGCAGATCTGTGTGTCAACAGGGAATTCAGAGGCTTAATAGGAGTGCAGCAGCAGCT    | 2100 |
| Db | 2041 | CAGCAGATCTGTGTGTCAACAGGGAATTCAGAGGCTTAATAGGAGTGCAGCAGCAGCT    | 2100 |
| QY | 2101 | CAGAGTCAAGAACTGATTAATTAACCAACAAGAACTCCGGTGCAGAACGATGATGTTTCC  | 2160 |
| Db | 2101 | CAGAGTCAAGAACTGATTAATTAACCAACAAGAACTCCGGTGCAGAACGATGATGTTTCC  | 2160 |
| QY | 2161 | TACCCAAACAATGTTCTTTATCAGGTGCATGACCCAGGGTTCTCAAGATCTGCCAG      | 2220 |
| Db | 2161 | TACCCAAACAATGTTCTTTATCAGGTGCATGACCCAGGGTTCTCAAGATCTGCCAG      | 2220 |
| QY | 2221 | CAGTCATACCAACAGCCAAATCATCTACTCTTACACAGGCAAGTCAAGGTCATCTCCAGCC | 2280 |
| Db | 2221 | CAGTCATACCAACAGCCAAATCATCTACTCTTACACAGGCAAGTCAAGGTCATCTCCAGCC | 2280 |
| QY | 2281 | ACTGGAATGCTGTTTATCTGTATGTACACCGCCACCCCTCAGAACCTTTAGGCTG       | 2340 |
| Db | 2281 | ACTGGAATGCTGTTTATCTGTATGTACACCGCCACCCCTCAGAACCTTTAGGCTG       | 2340 |
| QY | 2341 | ATTGGCCCACTGCCCCCTCAGACACTGTCCCAATATGTCACTAGTCAAGAACAAAC      | 2400 |



```

Db 179 VSYIPGQYRNS--NQQYRPLSHFVAYSPQRGQQLQPSQQLQPMMPQQAAYGATIG 237
Qy 636 VQGPQSQNVINNQQ--GTPVQSVWVSYPITNSSYQVPTNSQGLPQGSYQOPIMLP-N 751
Db 238 VQQ--PQNGGLASSQSSMGGQMGQGLVQYTPLSYQVPGVSDSQNVVQPPFQQLPLVPVS 296
Qy 752 QAGQSLPATGMPYCNTPPTPON 776
Db 297 QSVQGLPAPAGVPPYVYSMIPPAQN 321

```

## RESULT 2

```

US-10-449-902-41298
; Sequence 41298, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agricultural Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 41298
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-41298

```

Query Match 7.5%; Score 318.5; DB 6; Length 875;

Best Local Similarity 23.1%; Pred. No. 6,2e-08; Indels 319; Gaps 47;

Matches 207; Conservative 95; Mismatches 276; Indels 319; Gaps 47;

```

Qy 8 NOAIEGSGTEQETAPENGIVKSESLEDEEKLQRLLEAQNORRKSAGK----- 62
Db 109 SRTVLTSGGLQPTSTDEGTLLS-----SSDDGKPSLD 144
Qy 63 GKLTFS--LAVCESSAPPGGESLQDOESHLQUSFSSLOEEDKSRKDDSEKEDK 119
Db 145 GKSVAAGTTFALDEKESLRP-----DSSASV-----QASAEED--DL 179
Qy 120 NKDKTSEKPKIRMLSKDCSQEYTDSTGIDLHEF----- 152
Db 180 CYPPTGVPTSRMGSEDDGVAFRD--QLHEIHAMEPRRPPQAGPANGALPQVLVP 235
Qy 153 -----LINTLKNSRDMILLKMEQEIIFIDANN 182
Db 226 PEGPVGSIIPRSRQALATGGAAYAPDTKLEALE-NPRDIWVLEQDVIVDKAK 294
Qy 183 NHYKFPQMSYQRMVHRVAAVFGLDHNVDTGKSVIINKTSSTRIPE-----QRPC 235
Db 295 ESSLNLPOCNSFYRMIAHADIYVMLGSHSVDDSNSAVRLFRTPHCRIPQPLTGVTPSTA 354
Qy 236 EHLKDEKEESQKRIILKRDNSIDKDNQSVCSQESLFEVNSRLLEDNINENYKQRQ 295
Db 335 ASTPPSGPQIK--LKR--GID--PGPAIANGSNV--NSKTNDDSDSD--KKRP 401
Qy 296 LFRGND-----GSGRTSGSRQSSSENEKMSDHPORAMSTSDSSNRMLKPM 344
Db 402 VTBREARAYEARLRIMGSAKPTGSETSKED-----DERSAAGKSK 447
Qy 345 TKTAAPGGLIVLTRDSTSTSTGKLSKAGSESSSSAGSSGLSRTHPPLQSTPLVSGV 404
Db 448 KK-----QRNDSDDGFEARSAYSSYFAPSSSAGYT-----SQVAYONPQTV 492

```

```

Qy 405 AAGSPGVPEPENGIGGVAPSSSTSYLLPLEAATGIPGSLILNPH-TGQFVNPDDTP 463
Db 493 AQSFS--PY-----GNDDMAASPY-----GVQNO-----NSHSGSYTSGESG 533
Qy 464 AINPP-----TSQQLRS--AMV-----GQS-----QQQPPQQ 491
Db 534 SMYQVPQYDLSKQFNQMSFQNSMADQQQPPVYSFYAQQVAAQSGYGTSSAMQHQHQ 593
Qy 492 PSPPQQQVQ--PPQPMAG-----PLVT-QSVQGLQASQSQYQYAVSRP 534
Db 594 QSOHQOQOQOQYNAATQPTSTGFGPSPNFSDNNQMPVYTSQOQOQOQLHQOQOQNPFGQLP 653
Qy 535 PQ-----HLTPYS-PTQHPFMRDV-----ATQFGQMTLSRQSGETPE-PSS 575
Db 654 SQNLGRVPSALEHPLGWSKSPHRRPOSQTPVPGPIGNSNSSFPPQAPATGNSNNGHQ 713
Qy 576 GPVYPSLMPQPAQOPSYVIATGQQL--PTGFGSGGP--ISQOVLQPPSPQGFVQ 631
Db 714 G--VPYSLERQMSQ-----SQGSAPGSPHGHMGAGGPRMPSQPLTHLPQPV-FPRQ 764
Qy 632 PPP-----AQPVYVYPSGQYPTSTTQO-----YRMAPVQ- 662
Db 765 PSPVVALPPKPISGHQPLEH--ARQSPNVHQMSPOSSSIKMGAPASLEPKPPSPV 822
Qy 663 YNAOR--SQOMPQAQAGYQPVLSGQGFQGLIGVQOPQSQNVINNQGTPVQS 716
Db 823 FDASRFTPSQTPSTFNNAAARLRGGGLSGF-----GPMQPLNSSGMANG-NQSPRS 874

```

## RESULT 3

```

US-10-449-902-41078
; Sequence 41078, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agricultural Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 41078
; LENGTH: 1034
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-41078

```

Query Match 5.5%; Score 232.5; DB 6; Length 1034;

Best Local Similarity 20.1%; Pred. No. 0.0007; Indels 309; Gaps 43;

Matches 193; Conservative 99; Mismatches 357; Indels 309; Gaps 43;

```

Qy 2 SEQDLNOALA--EEGTEQETAPENGIVKSESLEDEEKLQRLLEAQNORRKSAG 58
Db 158 TEGELGVALASAOQIPADAKDAQPOLGALKSSILIAKDESIOAEVVRQODAKEMR 217
Qy 59 GA-----GKGLTR-----SLAVCESSAR-----PG 81
Db 218 EAREKEAFTTETISKAFDPVSGRALQKMSDDPLALDDHSSLRPAMASPSDAKASGA 277
Qy 82 ESLDOES-----IHLQSSPSSLOEE-----DKSRKDDSEKEDK----- 119
Db 278 GLEDSNKNKPMQDPAANEARFGIKSNYEENLYTKLDRTGKDFDRERADRIRAREWG 337
Qy 120 ---NKDKTSEKPKIRMLSKDCSQEYTDSTGIDLHEFLINTLKNSRDMILLKMEQEI 176
Db 338 QATNPPHMAER-----RGHADDSGVN-----BEDYKG 364

```



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

## OM protein - protein search, using sw model

Run on: July 5, 2006, 09:49:08 ; Search time 31 Seconds

(Without alignments)  
2295.561 Million cell updates/sec

Title: US-10-509-950-1

Sequence: 1 MSEGQDLMQAIABEGSTGEF.....SASCRTNCASMSNAGVQVRF 813

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database:

Issued Patents AA:\*

- 1: /EMC\_Celerra\_SIDS3/ptocata/2/1aa/5-COMB.pep:\*
- 2: /EMC\_Celerra\_SIDS3/ptocata/2/1aa/6-COMB.pep:\*
- 3: /EMC\_Celerra\_SIDS3/ptocata/2/1aa/7-COMB.pep:\*
- 4: /EMC\_Celerra\_SIDS3/ptocata/2/1aa/H-COMB.pep:\*
- 5: /EMC\_Celerra\_SIDS3/ptocata/2/1aa/PCRTUS-COMB.pep:\*
- 6: /EMC\_Celerra\_SIDS3/ptocata/2/1aa/RE-COMB.pep:\*
- 7: /EMC\_Celerra\_SIDS3/ptocata/2/1aa/backfilest.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description                           |
|------------|-------|-------------|--------|-------|---------------------------------------|
| 1          | 1510  | 35.7        | 971    | 2     | US-09-538-092-1332 Sequence 1332, Ap  |
| 2          | 473   | 11.2        | 107    | 2     | US-10-104-047-3736 Sequence 3736, Ap  |
| 3          | 293   | 6.9         | 1444   | 2     | US-09-949-016-9652 Sequence 9652, Ap  |
| 4          | 293   | 6.9         | 2004   | 2     | US-09-538-092-1371 Sequence 1371, Ap  |
| 5          | 293   | 6.9         | 2004   | 2     | US-09-949-016-6756 Sequence 6756, Ap  |
| 6          | 250   | 5.9         | 1185   | 2     | US-09-041-886-23 Sequence 23, Appl    |
| 7          | 250   | 5.9         | 1185   | 2     | US-09-538-092-1209 Sequence 1209, Ap  |
| 8          | 246.5 | 5.9         | 1184   | 2     | US-09-266-225D-18 Sequence 18, Appl   |
| 9          | 243.5 | 5.7         | 131    | 1     | US-08-991-300-2 Sequence 2, Appl      |
| 10         | 241   | 5.7         | 131    | 2     | US-09-270-767-44309 Sequence 44309, A |
| 11         | 233.5 | 5.5         | 2185   | 2     | US-09-854-856-36 Sequence 36, Appl    |
| 12         | 233.5 | 5.5         | 2185   | 2     | US-10-010-720-36 Sequence 36, Appl    |
| 13         | 233.5 | 5.5         | 2245   | 2     | US-09-854-856-4 Sequence 4, Appl      |
| 14         | 233.5 | 5.5         | 2245   | 2     | US-10-010-720-4 Sequence 4, Appl      |
| 15         | 233.5 | 5.5         | 2322   | 2     | US-09-854-856-34 Sequence 34, Appl    |
| 16         | 233.5 | 5.5         | 2322   | 2     | US-10-010-720-34 Sequence 34, Appl    |
| 17         | 233.5 | 5.5         | 2382   | 2     | US-09-854-856-2 Sequence 2, Appl      |
| 18         | 233.5 | 5.5         | 2382   | 2     | US-10-010-720-2 Sequence 2, Appl      |
| 19         | 232   | 5.5         | 2442   | 2     | US-09-514-247A-10 Sequence 10, Appl   |
| 20         | 232   | 5.5         | 2442   | 2     | US-09-538-092-1370 Sequence 1370, Ap  |
| 21         | 229   | 5.4         | 2414   | 1     | US-08-227-536-2 Sequence 2, Appl      |
| 22         | 229   | 5.4         | 2414   | 1     | US-09-538-092-1289 Sequence 1289, Ap  |
| 23         | 229   | 5.4         | 2414   | 5     | PCR-US95-04682-2 Sequence 2, Appl     |
| 24         | 227   | 5.4         | 1345   | 3     | US-10-433-794-17 Sequence 17, Appl    |
| 25         | 227   | 5.4         | 579    | 2     | US-09-668-119-3 Sequence 3, Appl      |
| 26         | 224.5 | 5.3         | 922    | 2     | US-09-854-856-38 Sequence 38, Appl    |

|    |       |     |      |   |  |
|----|-------|-----|------|---|--|
| 27 | 224.5 | 5.3 | 922  | 2 | US-10-010-720-38 Sequence 38, Appl     |
| 28 | 224.5 | 5.3 | 982  | 2 | US-09-854-856-6 Sequence 6, Appl       |
| 29 | 224.5 | 5.3 | 982  | 2 | US-10-010-720-6 Sequence 6, Appl       |
| 30 | 222   | 5.2 | 2441 | 1 | US-08-194-468-2 Sequence 2, Appl       |
| 31 | 222   | 5.2 | 2441 | 2 | US-08-361-739-2 Sequence 2, Appl       |
| 32 | 222   | 5.2 | 2441 | 2 | US-09-514-247A-8 Sequence 8, Appl      |
| 33 | 222   | 5.2 | 2441 | 2 | US-09-866-316-2 Sequence 2, Appl       |
| 34 | 221.5 | 5.2 | 802  | 2 | US-09-823-240A-2 Sequence 2, Appl      |
| 35 | 221.5 | 5.2 | 2079 | 2 | US-09-949-016-8301 Sequence 8301, Ap   |
| 36 | 219.5 | 5.2 | 1283 | 2 | US-09-489-039A-12097 Sequence 12097, A |
| 37 | 219   | 5.2 | 1326 | 2 | US-09-688-1888-15 Sequence 15, Appl    |
| 38 | 219   | 5.2 | 1326 | 2 | US-09-291-417D-15 Sequence 15, Appl    |
| 39 | 215   | 5.1 | 1274 | 2 | US-09-095-443-2 Sequence 2, Appl       |
| 40 | 215   | 5.1 | 1636 | 3 | US-10-433-794-1 Sequence 1, Appl       |
| 41 | 213   | 5.0 | 1179 | 2 | US-09-949-016-7088 Sequence 7088, Ap   |
| 42 | 212   | 5.0 | 915  | 2 | US-10-029-180-74 Sequence 74, Appl     |
| 43 | 211.5 | 5.0 | 2142 | 2 | US-09-538-092-1142 Sequence 1142, Ap   |
| 44 | 211.5 | 5.0 | 2142 | 2 | US-09-949-002-371 Sequence 371, Ap     |
| 45 | 211.5 | 5.0 | 2453 | 2 | US-09-632-033B-4 Sequence 4, Appl      |

## ALIGNMENTS

```
RESULT 1
US-09-538-092-1332
Sequence 1332, Application US/09538092
Patent No. 6753314
GENERAL INFORMATION:
APPLICANT: Glot, Loic
APPLICANT: Mansfield, Traci A.
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
FILE REFERENCE: 15966-542
CURRENT APPLICATION NUMBER: US/09/538, 092
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 60/127,352
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/178,965
PRIOR FILING DATE: 2000-02-01
NUMBER OF SEQ ID NOS: 1387
SOFTWARE: CurafastSeqformat Version 0.9
SEQ ID NO 1332
LENGTH: 971
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (0)...(0)
OTHER INFORMATION: Polypeptide Accession Number Q15032
US-09-538-092-1332
Query Match 35.7%; Score 1510; DB 2; Length 971;
Beet Local Similarity 40.6%; Pred. No. 1.5e-99;
Matches 332; Conservative 124; Mismatches 135; Indels 226; Gaps 23;
18 EOEETAPNGVYSPSLSEBKELEORLEAONERRKSKGAGKGLTRSLAVEESS 76
10 KDETAT-----MKDLAEAVKOTTRVENLKSEN-----GKLVKONHEHTEN 52
77 ARPGESELODESINLOUSFSFSLQEBDKSRKDSEREKDKNKTSEKFRIMLSKO 136
53 -----NIDLEKIQIQLT--QSPFEKPKPSDKBAEKKASDK-----LKR-KMLSRD 96
137 CSQETDSTGIDHFLINTLKNSRDMILKMEQETIDFLADNNNHKKFPQSSYOR 196
97 SSQETDSTGIDHFLVNTLKNSRDMILKMEQETIDFLADNNNHKKFPQSSYOR 156
197 MLVHRVAAYFGLDHVDQTKGKSVIINKTSSTRIPQRCFCEHLKDKGSESRFTIKRDN 256
157 MLVHRVAAYFGLDHVDQTKGKSVIINKTSSTRIPQRCFCEHLKDKGSESRFTIKRDN 216
257 SSIDKEDNQ-----SVCSQESLFLVENSRLLEDS 284
```



Db 217 S5FDKDNQVRIRLKDRRSKSIIEEREYORADRIFSDSLSCQENYIID--KRLQDE 274  
Qy 285 NINETYKROQLFRGNDRSGRTSGROSSENELKXSDHQRAMSDSDSSNNKLEPM 344  
Db 275 D-ASSTQQRQIRRVNADAGSTRNSQOSTENELKISE-PRPWSSTDSOSLRLNLEPAV 332  
Qy 345 TKTASFGITVLTTRGDSTSTRSTKLSKAGSESSSAGSSGLSRTHPLOSTPLVSGV 404  
Db 333 TKASSTGSLVLTGRGSSSSKSLGRISK----- 361  
Qy 405 AAGSPGCVPRPENGIGQVAPSTSTYLLPLEAATGIPPSILINPTGPPFVNDGTPA 464  
Db 362 -----TGQPFINPDGSPV 374  
Qy 465 IYNPTSGOPLRSAMVGSQQQPPQQPSPQPPQQVPPQPPQAGLVTVSGVGLQASSQ 524  
Db 375 VNPPTQQQVRSQVPE-----PPQPLPAPPPQ----- 403  
Qy 525 SVQPAVSPRPHLLPVSPTHFPRMRDVAOTFGOMTSLRSGSGETPRPSPGVPSLSM 584  
Db 404 -----PAAN-----HSQODNLSGSQFSHMSLAPQPSADGSDPHAA-MFOSTVV 446  
Qy 585 PPRAPQPSYIATST-----GQQLPTGFGSGSGPPIISQVLOPPPS 624  
Db 447 LQSPPOSGYIMTAAPRPHPRPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 502  
Qy 625 POGFVQOPRPAOMPVYV 664  
Db 503 QQQYIQQPSPF-QMPACTCAGHY-HSSQPPQYRPPVPSVHNSLNPQLPQPAQQTGVVIR 560  
Qy 685 SGGQGFQGLIGVQOPRPOSQVNVNNQO---GTPVQSVVWVSPYVMSYVQVPMTQSGGLPQ 741  
Db 561 NQOQNYQGVIGVQO-PQSGSLVSGQPSISNQLQGVVIPYTSVPTYQVSLPQSGQGIPIHQ 619  
Qy 742 SYQOPMLPNQAGGSLPATGMVYCVNTPPTPNL 778  
Db 620 TYQQPVMFPQSNQSGMPTTGMVYVSVIPRQOQNNL 656

RESULT 2  
US-10-104-047-3736  
; Sequence 3736, Application US/10104047  
; Patent No. 6943241  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. 6943241el full length cDNA  
; FILE REFERENCE: H1-A0105  
; CURRENT APPLICATION NUMBER: US/10/104,047  
; PRIOR FILING DATE: 2002-03-25  
; PRIOR APPLICATION NUMBER:  
; NUMBER OF SEQ ID NOS: 4096  
; SOFTWARE: PatentIn Ver. 2.1  
; LENGTH: 107  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-104-047-3736

Query Match 11.2%; Score 473; DB 2; Length 107;  
Best Local Similarity 86.5%; Pred. No. 1.5e-26;  
Matches 90; Conservative 2; Mismatches 4; Indels 8; Gaps 1;

Qy 718 MYSYPMSSV-----GVPMQSGQGLPQSYQOPIMLPNAGGSLPATGMVYCVN 769  
Db 4 IVQTKIVSVYLACVFLQVPMQSGQGLPQSYQOPIMLPNAGGSLPATGMVYCVN 63

Qy 770 TPTPTNNLRLIGPHCPSSSTVPVMSASCTNCASMSAGVQVKE 813  
Db 64 TPTPTNNLRLIGPHCPSSSTVPVMSASCTNCASMSAGVQVKE 107

RESULT 3

US-09-949-016-9652  
; Sequence 9652, Application US/09949016  
; Patent No. 6812319  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9652  
; LENGTH: 1444  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-9652

Query Match 6.9%; Score 293; DB 2; Length 1444;  
Best Local Similarity 23.4%; Pred. No. 5.2e-12;  
Matches 206; Conservative 102; Mismatches 359; Indels 214; Gaps 39;

Qy 9 QALAEFG-GRQETA-----TPENGIVKSESLEDEEK--LEIQRLEAQ 49  
Db 670 QAEAEAEAEEDDAASSEVPAPSPADSSNSPETETKEBEVEEKEKPRVSEEOQSEEE 729  
Qy 50 NERRRSKSGAKKXTRSLAVCESSARPGES---LQDSHILQUSFSLQEDK 105  
Db 730 QQLSEPEPEE-----BEDAALETKQNDHDADDEDGHLSETKKEL--BEQ 776  
Qy 106 SKRQDSE-----REKEDKNK---DKTSEKPIRMLSKDSQOEYDSTG 146  
Db 777 TREDVKEBPQVESFLDAMQKSRKIKQKEBELDSEBEP-----SHDTSVVSQEMAG 831  
Qy 147 I-DHEFLINTLKNSRDMILLKMEQELIDFLADNNHYKKFPQMSSTQRMVHRVAY 205  
Db 832 SEDDE-----EDSHKELEIEKEEELIPHELD-----LETQAVQSLTQEESEH 879  
Qy 206 PGLDHNVDOT--GKSYINKTSSTRIPRQFCE--HLKDEKGEESQKRLFKDNSIDK 261  
Db 880 EAYQDCETTLAQCOTLOSITQADEDPQMSWVEDCHASHEHS-----PISSVQS 928  
Qy 262 EDNQSQSGESLFEVNSRLLEDNSICNETYKKROLFRGNRDGSGRT-----SGSROSSSE 316  
Db 929 HPSQSVRSVS---PRVPALE-----SGYQIISPEQSLAPSM 964  
Qy 317 NELKNSDHQRAMSTDS-----DSSNRNLKPMATKTASFGITVLTTRGDSTSTRSTGKL 371  
Db 965 QNMETSPPMADVPVSVDHSCQVVDGSGFSDLSGISTETENY-----NPPSSYDSTMG 1015  
Qy 372 SKAGSESSSAGSSGSLSTRHPLOSTPLVS-GVAGSGCVPRPENGIGQVAPSTSY 430  
Db 1016 STCGSSSSQSSCSYGLSSSSSLTQSSCVVTOQMMSGSSCSMMQOS---SQPAMNS 1071  
Qy 431 ILPLEAATGIPGSLILNPTGOFVNPDPATYVNPPTSQPLRSAMVGSQQQPPQ 490  
Db 1072 IKSQSCVERPFSN-----QQQPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 1121  
Qy 491 QPSPPQPPQVQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 1168  
Db 1122 QPPQPPQPP 1168  
Qy 547 FPMRDVATQFGOMTIS-----ROSGETPRPSPGVPSLSMPPQPAQGSYV 595  
Db 1169 ISYERIRGDFGAGSYSQPSATFSLAKLQQLINTIMDPLAMYSNHPATYSTATVSLN 1228  
Qy 596 ASTGQQLPTGFGSGSGPPIISQVLOPPSPQGFVQO--PPPAO-----MPVYVPS 644

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: July 5, 2006, 09:47:58 ; Search time 309 Seconds

(without alignments)  
2433.779 Million cell updates/sec

Title: US-10-509-950-1

Perfect score: 4235  
Sequence: 1 MSBQGLNQAIAEAGTEOE.....SASCRTNCASMSNAGQVKE 813

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: uniprot:2.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID          | Description        |
|------------|--------|-------------|--------|----------------|--------------------|
| 1          | 411.1  | 97.1        | 793    | 2 Q4G0V4 HUMAN | Q4G0V4 homo sapien |
| 2          | 3997.5 | 94.4        | 812    | 2 Q86VJ1 HUMAN | Q86VJ1 homo sapien |
| 3          | 3641   | 86.0        | 809    | 2 Q7TS83 MOUSE | Q7TS83 mus musculu |
| 4          | 3533   | 83.4        | 789    | 2 Q8C038 MOUSE | Q8C038 mus musculu |
| 5          | 3433.5 | 81.1        | 771    | 2 Q8C018 MOUSE | Q8C018 mus musculu |
| 6          | 3421   | 80.8        | 807    | 2 Q91Y59 MOUSE | Q91Y59 mus musculu |
| 7          | 3414.5 | 80.6        | 651    | 2 Q9UF93 HUMAN | Q9UF93 homo sapien |
| 8          | 2450   | 57.9        | 601    | 2 Q3U0U8 MOUSE | Q3U0U8 mus musculu |
| 9          | 1831.5 | 43.2        | 784    | 2 Q53S01 HUMAN | Q53S01 homo sapien |
| 10         | 1831.5 | 43.2        | 1099   | 1 R3HD1 HUMAN  | 015032 homo sapien |
| 11         | 1827.5 | 43.2        | 1061   | 2 Q5R7K5 PONY  | Q5R7K5 pongo pygma |
| 12         | 1789   | 42.2        | 335    | 2 Q49A56 HUMAN | Q49A56 homo sapien |
| 13         | 1570   | 37.1        | 1044   | 1 R3HD2 MOUSE  | Q80t66 mus musculu |
| 14         | 1493   | 35.3        | 962    | 1 R3HD2 HUMAN  | Q9Y2K5 homo sapien |
| 15         | 1463   | 34.5        | 1009   | 2 Q505L6 XENLA | Q505L6 xenopus lae |
| 16         | 1316.5 | 31.1        | 999    | 2 Q4SS78 TETNG | Q4SS78 tetraodon n |
| 17         | 1298   | 30.6        | 823    | 2 Q5R8U0 PONY  | Q5R8U0 pongo pygma |
| 18         | 1296.5 | 30.6        | 549    | 2 Q80ZB9 MOUSE | Q80ZB9 mus musculu |
| 19         | 1231   | 29.1        | 372    | 2 Q8C0Y6 MOUSE | Q8C0Y6 mus musculu |
| 20         | 1109.5 | 26.2        | 466    | 2 Q6NRG3 XENLA | Q6NRG3 xenopus lae |
| 21         | 1075.5 | 25.4        | 1066   | 2 Q4SEC3 TETNG | Q4SEC3 tetraodon n |
| 22         | 1036.5 | 24.5        | 248    | 2 Q8BMM1 MOUSE | Q8BMM1 mus musculu |
| 23         | 1030.5 | 24.3        | 228    | 2 Q8BUD1 MOUSE | Q8BUD1 mus musculu |
| 24         | 976.5  | 23.1        | 455    | 2 Q5PPF9 RAT   | Q5PPF9 rattus norv |
| 25         | 942.5  | 22.3        | 352    | 2 Q4R3F1 MACRA | Q4R3F1 macaca fasc |
| 26         | 741    | 17.5        | 347    | 2 Q4RNM3 TETNG | Q4RNM3 tetraodon n |
| 27         | 736.5  | 17.4        | 681    | 2 Q3ZCT5 HUMAN | Q3ZCT5 homo sapien |
| 28         | 679    | 16.0        | 657    | 2 Q5XKA9 MOUSE | Q5XKA9 mus musculu |
| 29         | 672    | 15.9        | 656    | 2 Q2MT9 HUMAN  | Q2MT9 homo sapien  |
| 30         | 652    | 15.4        | 445    | 2 Q9SKJ7 MACRA | Q9SKJ7 macaca fasc |
| 31         | 617    | 14.6        | 239    | 2 Q3URW7 MOUSE | Q3URW7 mus musculu |

|    |       |      |      |                |                    |
|----|-------|------|------|----------------|--------------------|
| 32 | 430.5 | 10.2 | 108  | 2 Q6NYC3 HUMAN | Q6NYC3 homo sapien |
| 33 | 430   | 10.2 | 89   | 1 ARP21 HUMAN  | Q9UB10 homo sapien |
| 34 | 430   | 10.2 | 89   | 1 ARP21 PONY   | Q5R6X9 pongo pygma |
| 35 | 430   | 10.2 | 89   | 2 Q49AK3 HUMAN | Q49AK3 homo sapien |
| 36 | 430   | 10.2 | 89   | 2 Q4R4W2 MACRA | Q4R4W2 macaca fasc |
| 37 | 429   | 10.1 | 1738 | 2 Q2M075 DROPS | Q2M075 drosophila  |
| 38 | 414   | 9.8  | 1823 | 2 Q5BIH9 DROME | Q5BIH9 drosophila  |
| 39 | 414   | 9.8  | 1823 | 2 Q7KV61 DROME | Q7KV61 drosophila  |
| 40 | 410.5 | 9.7  | 1818 | 2 ENC DROME    | Q6MSX1 drosophila  |
| 41 | 396   | 9.4  | 89   | 2 Q7M2N1 BOVIN | Q7M2N1 bos taurus  |
| 42 | 355.5 | 8.4  | 88   | 2 Q5FV10 RAT   | Q5FV10 rattus norv |
| 43 | 344.5 | 8.1  | 88   | 1 ARP21 MOUSE  | Q9DCB4 mus musculu |
| 44 | 341   | 8.1  | 558  | 2 Q9NAP3 CAEBL | Q9NAP3 caenorhabdi |
| 45 | 318   | 7.5  | 1080 | 2 Q559H7 DICDI | Q559H7 dictyosteli |

#### ALIGNMENTS

RESULT 1  
Q4G0V4 HUMAN PRELIMINARY; PRT; 793 AA.  
ID Q4G0V4 HUMAN  
AC Q4G0V4  
DT 30-AUG-2005, integrated into UniProtKB/TREMBL.  
DT 30-AUG-2005, sequence version 1.  
DT 07-FEB-2006, entry version 2.  
DE ARP-21 protein.  
GN Name=ARP-21;  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TTSUB=Brain;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klisner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.T., Wang J., Hsieh F.,  
RA Datchenko U., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Uddin T.B., Tothiyuki S., Carninci P., Prange C.,  
RA Raba S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
RA Scheraga A., Schein U.E., Jones S.J.M., Marra W.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[2]  
RP NUCLEOTIDE SEQUENCE.  
RC TTSUB=Brain;  
RG NIH MGC Project;  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
distributed under the Creative Commons Attribution-NonCommercial license  
CC EMBL, BC036399; AH36399.1; -, mRNA.  
DR SMR; Q4G0V4; 147-257.  
DR GO; GO:0003676; F:uncle acid binding; IEA.  
DR InterPro; IPR001374; R3H\_ss\_bd.  
DR Pfam; PF01424; R3H; 1.  
DR SMART; SM00393; R3H; 1.  
DR PROSITE; PS1061; R3H; 1.

SQ SEQUENCE 793 AA; 86131 MW; F523AEF3E39114BE CRC64;  
 Query Match 97.1%; Score 4111; DB 2; Length 793;  
 Best Local Similarity 97.5%; Pred. No. 1,8e-162;  
 Matches 793; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

```

1 MSBQGLNDAIAEAGTEGTETATPENGIVKSESLDEBEKLELORLEAQNQERRKSKSGA 60
1 MSBQGLNDAIAEAGTEGTETATPENGIVKSESLDEBEKLELORLEAQNQERRKSKSGA 60
61 GKGKLTSLAVCESSARPGESLQDESIIHLQSSFSLSQEDSKRKDSEREKEDKN 120
61 GKGKLTSLAVCESSARPGESLQDESIIHLQSSFSLSQEDSKRKDSEREKEDKN 120
121 KDKTSKPKIRMLSKDCSQEYTDSTGIDLHEFLINLKNNRPRMILLKKEOEIIIDFIAD 180
121 KDKTSKPKIRMLSKDCSQEYTDSTGIDLHEFLINLKNNRPRMILLKKEOEIIIDFIAD 180
121 KDKTSKPKIRMLSKDCSQEYTDSTGIDLHEFLINLKNNRPRMILLKKEOEIIIDFIAD 180
181 NNNHYKFPOMSSYQMLVHRVAAYFGLDHNVDQTKSVIINTKSTRIPEDQFCEHLKD 240
181 NNNHYKFPOMSSYQMLVHRVAAYFGLDHNVDQTKSVIINTKSTRIPEDQFCEHLKD 240
241 EKGEESQKRFILKRDNSIDKEDNOSVCSQESLFEVNSRLLEDNSI CNETTKRQLFRGN 300
241 EKGEESQKRFILKRDNSIDKEDNOSVCSQESLFEVNSRLLEDNSI CNETTKRQLFRGN 300
301 RGSQGTSSGROSSSENELKMSDHQAMSTDBDSSNRMLKPMNTKTAASGCTIVLTRGD 360
301 RGSQGTSSGROSSSENELKMSDHQAMSTDBDSSNRMLKPMNTKTAASGCTIVLTRGD 360
361 STSSTSTGKLSTAGSESSSSAGSSSLSTRHPLOSTPLVSGVAAGSPGVYPENGIG 420
361 STSSTSTGKLSTAGSESSSSAGSSSLSTRHPLOSTPLVSGVAAGSPGVYPENGIG 420
421 GGVAPSTSTYIILLPLEAATGIPPGSILLNPHTCQPFVNPDPGTAIYNPPTSOQPLRSAMY 480
421 GGVAPSTSTYIILLPLEAATGIPPGSILLNPHTCQPFVNPDPGTAIYNPPTSOQPLRSAMY 480
481 GSGQQQPPQQQPPQQQPPQQQPPQQQPPQQQPPQQQPPQQQPPQQQPPQQQPPQQ 540
481 GSGQQQPPQQQPPQQQPPQQQPPQQQPPQQQPPQQQPPQQQPPQQQPPQQQPPQQ 540
541 VSETFQHPKRDVATQFGQWTLSSROSSGSETPEPSPGVVPSLMPPOAQPSYVIASTGQ 600
541 VSETFQHPKRDVATQFGQWTLSSROSSGSETPEPSPGVVPSLMPPOAQPSYVIASTGQ 600
601 QLEPTGFGSGGPPISQOVLQPPSPGQFVQPPPAQMPVYVYPSGQYPTISTTQYRPMAP 660
601 QLEPTGFGSGGPPISQOVLQPPSPGQFVQPPPAQMPVYVYPSGQYPTISTTQYRPMAP 660
661 VQVNAARSOQMPAAQOAGYQPVLSGQGGQGLIGVQPPQSOQVILNNOGQTVQSMVS 720
661 VQVNAARSOQMPAAQOAGYQPVLSGQGGQGLIGVQPPQSOQVILNNOGQTVQSMVS 720
721 YPTMSYQVPMTGSGGGLPQOSYQOPIMLPNQAGGSLPATGMVYCNVTPTPTPNNLRL 780
721 YPTMSYQVPMTGSGGGLPQOSYQOPIMLPNQAGGSLPATGMVYCNVTPTPTPNNLRL 780
781 IGPHCPSSTVPMVASCRINCASMSNAGMVKF 813
781 IGPHCPSSTVPMVASCRINCASMSNAGMVKF 813
  
```

RESULT 2  
 ID 086V31 HUMAN PRELIMINARY; PRT; 812 AA.  
 AC 086V31;  
 DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.  
 DT 01-JUN-2003, sequence version 1.  
 DE 07-FEB-2006, entry version 13.  
 DE Cyclic AMP-regulated phosphoprotein, 21 kD, isoform 1.  
 GN Name=ARPP-21;  
 OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;  
 OC Homo.  
 NCBI\_TaxID=9606;  
 RN (1)  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Testis;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins P.S., Wagner K.H., Shemen C.M., Schler G.D.,  
 RA Altschul S.F., Zeeberg B., Butow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang Y., Hsieh F.,  
 RA Diachenko L., Matsuda K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Schetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Maita M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN (2)  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Testis;  
 RG NIH NIG Project;  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NonCommercial  
 CC License.  
 CC EMBL, BC051828; AAH51828.1; -; mRNA.  
 DR SMR; 086V31; 147-257.  
 DR Ensembl; ENSG00000172995; Homo sapiens.  
 DR GO; GO:0003676; F nucleic acid binding; IEA.  
 DR InterPro; IPR001374; R3H\_ss\_Bd.  
 DR Pfam; PF01424; R3H; 1.  
 DR SMART; SM00393; R3H; 1.  
 DR PROSITE; PS51061; R3H; 1.  
 SQ SEQUENCE 812 AA; 89196 MW; 73981A8341BB3359 CRC64;  
 Query Match 94.4%; Score 3997.5; DB 2; Length 812;  
 Best Local Similarity 91.9%; Pred. No. 9e-158;  
 Matches 778; Conservative 0; Mismatches 0; Indels 69; Gaps 2;

```

1 MSBQGLNDAIAEAGTEGTETATPENGIVKSESLDEBEKLELORLEAQNQERRKSKSGA 60
1 MSBQGLNDAIAEAGTEGTETATPENGIVKSESLDEBEKLELORLEAQNQERRKSKSGA 60
61 GKGKLTSLAVCESSARPGESLQDESIIHLQSSFSLSQEDSKRKDSEREKEDKN 120
61 GKGKLTSLAVCESSARPGESLQDESIIHLQSSFSLSQEDSKRKDSEREKEDKN 120
121 KDKTSKPKIRMLSKDCSQEYTDSTGIDLHEFLINLKNNRPRMILLKKEOEIIIDFIAD 180
121 KDKTSKPKIRMLSKDCSQEYTDSTGIDLHEFLINLKNNRPRMILLKKEOEIIIDFIAD 180
121 KDKTSKPKIRMLSKDCSQEYTDSTGIDLHEFLINLKNNRPRMILLKKEOEIIIDFIAD 180
181 NNNHYKFPOMSSYQMLVHRVAAYFGLDHNVDQTKSVIINTKSTRIPEDQFCEHLKD 240
181 NNNHYKFPOMSSYQMLVHRVAAYFGLDHNVDQTKSVIINTKSTRIPEDQFCEHLKD 240
241 EKGEESQKRFILKRDNSIDKEDNOSVCSQESLFEVNSRLLEDNSI CNETTKRQLFRGN 300
241 EKGEESQKRFILKRDNSIDKEDNOSVCSQESLFEVNSRLLEDNSI CNETTKRQLFRGN 300
301 RGSQGTSSGROSSSENELKMSDHQAMSTDBDSSNRMLKPMNTKTAASGCTIVLTRGD 360
301 RGSQGTSSGROSSSENELKMSDHQAMSTDBDSSNRMLKPMNTKTAASGCTIVLTRGD 360
361 STSSTSTGKLSTAGSESSSSAGSSSLSTRHPLOSTPLVSGVAAGSPGVYPENGIG 420
361 STSSTSTGKLSTAGSESSSSAGSSSLSTRHPLOSTPLVSGVAAGSPGVYPENGIG 420
421 GGVAPSTSTYIILLPLEAATGIPPGSILLNPHTCQPFVNPDPGTAIYNPPTSOQPLRSAMY 480
421 GGVAPSTSTYIILLPLEAATGIPPGSILLNPHTCQPFVNPDPGTAIYNPPTSOQPLRSAMY 480
481 GSGQQQPPQQQPPQQQPPQQQPPQQQPPQQQPPQQQPPQQQPPQQQPPQQQPPQQ 540
481 GSGQQQPPQQQPPQQQPPQQQPPQQQPPQQQPPQQQPPQQQPPQQQPPQQQPPQQ 540
541 VSETFQHPKRDVATQFGQWTLSSROSSGSETPEPSPGVVPSLMPPOAQPSYVIASTGQ 600
541 VSETFQHPKRDVATQFGQWTLSSROSSGSETPEPSPGVVPSLMPPOAQPSYVIASTGQ 600
601 QLEPTGFGSGGPPISQOVLQPPSPGQFVQPPPAQMPVYVYPSGQYPTISTTQYRPMAP 660
601 QLEPTGFGSGGPPISQOVLQPPSPGQFVQPPPAQMPVYVYPSGQYPTISTTQYRPMAP 660
661 VQVNAARSOQMPAAQOAGYQPVLSGQGGQGLIGVQPPQSOQVILNNOGQTVQSMVS 720
661 VQVNAARSOQMPAAQOAGYQPVLSGQGGQGLIGVQPPQSOQVILNNOGQTVQSMVS 720
721 YPTMSYQVPMTGSGGGLPQOSYQOPIMLPNQAGGSLPATGMVYCNVTPTPTPNNLRL 780
721 YPTMSYQVPMTGSGGGLPQOSYQOPIMLPNQAGGSLPATGMVYCNVTPTPTPNNLRL 780
781 IGPHCPSSTVPMVASCRINCASMSNAGMVKF 813
781 IGPHCPSSTVPMVASCRINCASMSNAGMVKF 813
  
```

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: July 5, 2006, 09:48:03 ; Search time 45 Seconds

(without alignments)  
1738.316 Million cell updates/sec

Title: US-10-509-950-1

Perfect score: 4235  
Sequence: 1 MSBQDNLNQAIAEBGTEQE.....SASCRTNCASMSNAGVXKF 813

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: pfr\_81:\*  
2: pfr1:\*  
3: pfr2:\*  
4: pfr3:\*  
5: pfr4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID  | Description           |
|------------|--------|-------------|--------|--------|-----------------------|
| 1          | 3414.5 | 80.5        | 651    | T42644 | hypothetical prote    |
| 2          | 293.5  | 9.4         | 89     | A34957 | CAMP-regulated pro    |
| 3          | 282.5  | 6.9         | 2715   | T13049 | eyelid - fruit fly    |
| 4          | 279    | 6.6         | 838    | EWTHW  | glutenin, high mol    |
| 5          | 274.5  | 6.5         | 848    | S02262 | glutenin, high mol    |
| 6          | 274.5  | 6.5         | 753    | JC2099 | glutenin, high mol    |
| 7          | 273.5  | 6.5         | 789    | A30843 | glutenin, high mol    |
| 8          | 267.5  | 6.3         | 791    | JN0690 | glutenin, high mol    |
| 9          | 265    | 6.3         | 1006   | T42731 | atrophin-1 related    |
| 10         | 262.5  | 6.2         | 1026   | T20369 | hypothetical prote    |
| 11         | 262.5  | 6.2         | 357    | S18236 | omega secalin prec    |
| 12         | 262    | 6.2         | 782    | T48722 | hypothetical prote    |
| 13         | 260.5  | 6.2         | 347    | T05737 | probable hordein C    |
| 14         | 258.5  | 6.1         | 660    | A24266 | glutenin high mole    |
| 15         | 258    | 6.1         | 3498   | T22330 | hypothetical prote    |
| 16         | 256.5  | 6.1         | 357    | S18235 | omega secalin prec    |
| 17         | 254    | 6.0         | 1119   | T50995 | related to cytochrome |
| 18         | 250    | 5.9         | 830    | S15720 | glutenin high mole    |
| 19         | 250    | 5.9         | 1184   | S50832 | atrophin-1 - human    |
| 20         | 249.5  | 5.9         | 1420   | T37781 | probable cytochrome   |
| 21         | 249    | 5.9         | 648    | S04832 | glutenin high mole    |
| 22         | 248.5  | 5.9         | 1184   | G01763 | atrophin-1 - human    |
| 23         | 248    | 5.9         | 748    | T04011 | hypothetical prote    |
| 24         | 248    | 5.9         | 815    | JN0689 | glutenin, high-mol    |
| 25         | 245.5  | 5.8         | 815    | B30843 | glutenin high mole    |
| 26         | 243.5  | 5.7         | 872    | S33015 | hypothetical prote    |
| 27         | 243.5  | 5.7         | 992    | A31666 | hypothetical prote    |
| 28         | 243    | 5.7         | 705    | S18733 | glutenin high mole    |
| 29         | 241.5  | 5.7         | 1234   | T30160 | hypothetical prote    |

|    |       |     |      |         |                     |
|----|-------|-----|------|---------|---------------------|
| 30 | 238.5 | 5.6 | 1088 | H96747  | unknown protein T1  |
| 31 | 237.5 | 5.6 | 1307 | T25563  | hypothetical prote  |
| 32 | 235   | 5.5 | 528  | T147141 | gastric mucin (clo  |
| 33 | 235   | 5.5 | 1494 | T14355  | protein-tyrosine-p  |
| 34 | 234.5 | 5.5 | 1018 | S44758  | Cl4B9.6 protein -   |
| 35 | 234   | 5.5 | 310  | T06211  | C-hordein precursor |
| 36 | 234   | 5.5 | 2422 | T12687  | ALR protein homolo  |
| 37 | 233.5 | 5.5 | 786  | T01456  | extensin homology F |
| 38 | 233   | 5.5 | 1076 | T24887  | hypothetical prote  |
| 39 | 231   | 5.5 | 802  | S48529  | NAB3 protein - yea  |
| 40 | 231   | 5.5 | 1074 | T24877  | hypothetical prote  |
| 41 | 230   | 5.4 | 620  | S06733  | hydroxyproline-ric  |
| 42 | 229   | 5.4 | 1611 | T38236  | hypothetical prote  |
| 43 | 229   | 5.4 | 2414 | A54277  | transcription adap  |
| 44 | 228.5 | 5.4 | 2129 | T27431  | hypothetical prote  |
| 45 | 226   | 5.3 | 374  | T05923  | glutenin low molec  |

## ALIGNMENTS

RESULT 1  
T42644  
hypothetical protein DKFZP566N1047.1 - human (fragment)  
C/Species: Homo sapiens (man)  
C/Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jul-2004  
C/Accession: T42644  
R/Ottensmeyer, B.; Obermayer, B.; Mewes, H.W.; Gassenhuber, U.; Wilmann, S.  
submitted to the Protein Sequence Database, November 1999  
A/Reference number: Z22231  
A/Accession: T42644  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-651 <AAs>  
A/Cross-references: UNIPROT:Q9UP93; UNIPARC:UPI000006F0B2; EMBL:AL133109  
A/Experimental source: fetal kidney; clone DKFZP566N1047  
C/Genetics:  
A/Note: DKFZP566N1047.1

| Query Match | Score | DB 2: Length   | Best Local Similarity | Pred. No. | Matches | Conservative | Mismatches | Indels | Gaps |
|-------------|-------|--|-----------------------|-----------|---------|--------------|------------|--------|------|
| QY          | 162   | RDMLILKMEOEIIFDIADNNHYKFPOMSSYQRMVLRVAAVFGIDHNVDTGKSVII  | 80.6%                 | 221       | 651     | 0            | 0          | 1      | 1    |
| DB          | 1     | RDMLILKMEOEIIFDIADNNHYKFPOMSSYQRMVLRVAAVFGIDHNVDTGKSVII  | 99.8%                 | 60        | 651     | 0            | 0          | 1      | 1    |
| QY          | 222   | NKTSSTRIPBQRFCEHLKDEKESQKRFILKRDNSIDKEDNQVCSQESLFEVNSRL  | 80.6%                 | 281       | 651     | 0            | 0          | 1      | 1    |
| DB          | 61    | NKTSSTRIPBQRFCEHLKDEKESQKRFILKRDNSIDKEDNQVCSQESLFEVNSRL  | 99.8%                 | 119       | 651     | 0            | 0          | 1      | 1    |
| QY          | 282   | EDSNICNERYKKRQIFRGRDSSGRTSGRSSSENELKMSDHRQVMSSTDSSNRNLK  | 80.6%                 | 341       | 651     | 0            | 0          | 1      | 1    |
| DB          | 130   | EDSNICNERYKKRQIFRGRDSSGRTSGRSSSENELKMSDHRQVMSSTDSSNRNLK  | 99.8%                 | 179       | 651     | 0            | 0          | 1      | 1    |
| QY          | 342   | PAMTKTASFGITVLTTRGDSSTSTRSTGLSAGSSSSSAGSSGLSTHPTPLV      | 80.6%                 | 401       | 651     | 0            | 0          | 1      | 1    |
| DB          | 180   | PAMTKTASFGITVLTTRGDSSTSTRSTGLSAGSSSSSAGSSGLSTHPTPLV      | 99.8%                 | 239       | 651     | 0            | 0          | 1      | 1    |
| QY          | 402   | SGVAASPGCVYPENGIQGVAPSTSYLLPLEAATGIPGSIILNHTGQPFVNPDG    | 80.6%                 | 461       | 651     | 0            | 0          | 1      | 1    |
| DB          | 240   | SGVAASPGCVYPENGIQGVAPSTSYLLPLEAATGIPGSIILNHTGQPFVNPDG    | 99.8%                 | 299       | 651     | 0            | 0          | 1      | 1    |
| QY          | 462   | TPAINTPTSOQPLASAWGSGQQPQQQBPQQQQVQPPQMAAGPLVQSVGLQA      | 80.6%                 | 521       | 651     | 0            | 0          | 1      | 1    |
| DB          | 300   | TPAINTPTSOQPLASAWGSGQQPQQQBPQQQQVQPPQMAAGPLVQSVGLQA      | 99.8%                 | 359       | 651     | 0            | 0          | 1      | 1    |
| QY          | 522   | SSQSVQYPAVSFPFPHLLPVSPFQHPMRDVAATGFGQMTLRSQSGTPEPPSGVYPS | 80.6%                 | 581       | 651     | 0            | 0          | 1      | 1    |
| DB          | 360   | SSQSVQYPAVSFPFPHLLPVSPFQHPMRDVAATGFGQMTLRSQSGTPEPPSGVYPS | 99.8%                 | 419       | 651     | 0            | 0          | 1      | 1    |
| QY          | 582   | SLMPQAPQPSYVIASTGQQLPTGFGSGGPPISQGVLOPPSPGQFVQPPAQNPVY   | 80.6%                 | 641       | 651     | 0            | 0          | 1      | 1    |
| DB          | 582   | SLMPQAPQPSYVIASTGQQLPTGFGSGGPPISQGVLOPPSPGQFVQPPAQNPVY   | 99.8%                 | 641       | 651     | 0            | 0          | 1      | 1    |

Db 420 SLMPDPAQPSYVIASTGQQLPTGSGSGSPISQVLPSPSGVQPPPAQMHVY 479  
QY 642 YPSGQVPTSTTQOYRPMAPVQVNAORSQOMPOAAQAGVQPVLSGGQGFGLGVQOQPQ 701  
Db 480 YPSGQVPTSTTQOYRPMAPVQVNAORSQOMPOAAQAGVQPVLSGGQGFGLGVQOQPQ 539  
QY 702 SONVINOOGTPTVQVWVSYPVMSYQVPMTOGSGGLPQOSYQOPIPLPNOAGSGSLPAT 761  
Db 540 SQVIVINNOGTPTVQVWVSYPVMSYQVPMTOGSGGLPQOSYQOPIPLPNOAGSGSLPAT 599  
QY 762 GMFVYCNVTPPTPQNNRLIGPHCPSTVPVMSASCTTCASMSNMGQVYF 813  
Db 600 GMFVYCNVTPPTPQNNRLIGPHCPSTVPVMSASCTTCASMSNMGQVYF 651

RESULT 2  
A34957  
CAMP-regulated phosphoprotein 21 - bovine  
N:Alternate names: ARPP-21  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 13-Jul-1990 #sequence\_revision 13-Jul-1990 #text\_change 09-Jul-2004  
C:Accession: A34957, A34956  
R:Kurthare, T.; Ehrlich, M.E.; Horiuchi, J.; Nasu, T.; Greengard, P.  
J. Neurosci. 9, 3638-3644, 1989  
A:Title: ARPP-21, a cyclic AMP-regulated phosphoprotein enriched in dopamine-innervated  
A:Reference number: A34957, NCID:90011251, PMID:2552037  
A:Accession: A34957  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-89 <KUR>  
A:Cross-references: UNIPROT:Q7M2N1; UNIPARC:UPI000017C4B9  
J. Williams, K.R.; Hemming Jr., H.C.; LoPresti, M.B.; Greengard, P.  
J. Neurosci. 9, 3631-3637, 1989  
A:Title: ARPP-21, a cyclic AMP-regulated phosphoprotein enriched in dopamine-innervated  
A:Reference number: A34956, NCID:90011250, PMID:2552036  
A:Accession: A34956  
A:Molecule type: protein  
A:Residues: 2-25, 'D', 27-89 <WIL>  
A:Cross-references: UNIPARC:UPI000017C4BA  
A:Note: this sequence was obtained from isoform ARPP-21B  
C:Keywords: acetylated amino end, phosphoprotein 21 #status predicted <MAT>  
F:2-89/Product: CAMP-regulated phosphoprotein 21 #status experimental  
F:2/Modified site: acetylated amino end (Ser) (in mature form) #status experimental  
F:56/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #status experim

Query Match 9.4%; Score 396; DB 2; Length 89;  
Best Local Similarity 90.8%; Pred. No. 3,4e-12;  
Matches 79; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSBQGLNOAIAEGGTGTEGTATPENGIVKESLDEBEKLEORLEAONQERRKSKSGA 60  
Db 1 MSBPGLSQTVIEGEGPEGTATPENGIVKESLDEBEKLEORLVANQERRKSKSGA 60  
QY 61 GKGLTRSLAVCESSARPGEESLQDQ 87  
Db 61 GKGLTRSLAVCESSARPGEESLQDQ 87

RESULT 3  
T13049  
eyelid - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C>Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: T13049  
R:Freeman, J.E.; Luk, A.; Rubin, G.M.; Heberlein, U.  
Submitted to the EMBL Data Library, March 1998  
A:Reference number: Z17592  
A:Accession: T13049  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-2715 <TRB>  
A:Cross-references: UNIPROT:Q9IN94; UNIPARC:UPI000007D87E; EMBL:AF053091; NCID:g2981220;  
C:Genetics:

A:Gene: eid  
A:Cross-references: FlyBase:FBgn0003013  
C:Function:  
A:Description: could act as a transcription factor antagonistic to the Wg pathway  
C:Keywords: DNA binding

Query Match 6.9%; Score 293.5; DB 2; Length 2715;  
Best Local Similarity 22.0%; Pred. No. 8,7e-06;  
Matches 215; Conservative 71; Mismatches 334; Indels 357; Gaps 45;

QY 4 QEDLNOAIAEGGTGTEGTATPENGIVKESLDEBEKLEORLEAONQERRKSKSGA 60  
Db 835 QGQVQPGQAGASGPEGGEHISODNGI-----SSS 865

QY 61 GKGLTRSLAVCESSARPGEESLQDQESIHQLSSFSLSQED-----KSRKDSERE 114  
Db 866 GPTGAAGMAHVAIVSVTTGPDGISM-DEVQOSTLSNASSAGBDPQCTTPKSKRND-y 922

QY 115 KKKDKNDKTSSEKPKIRMLSKDCSQEYTDSTGIDLHEFLINTLKNNSRDMILKMEQEI 174  
Db 923 SQSHLAPSTSPHPVVMHGGGGEYEDMSSP----- 954

QY 175 IDPLADNNHYKKFPQMSYQMLYHRVAAYFGLDHNVDQTKSVIINKTSRTRIPQR 234  
Db 955 -----FNWPRPAGSPQVFNH----- 977

QY 235 CEHLDEKGEESQKPILRDNSIDKEDNOSVCSQESLFEVNSR--LLEDNINCNETKY 292  
Db 978 RSTITTTKSDSLCKYEMDDN-----PDRGMLDKLRFMERRRPITACPTISKQPLD 1032

QY 293 KQOLFKNRNDGSGRTSGRSSENELKMSDQRAWSTDSNBNLKPAMTKTASFGG 352  
Db 1033 LYRLVLYVERGAGFVETKSKT-----WKDIAGLIGIGASSAAVTLRKHYTK----- 1080

QY 353 TVLVTL-----RGD-----STSSFRSGLSKASGE-----SSSAGSGSL 388  
Db 1081 -MLTFECHFDGDDIDPLPIQVEBAGSKKTAKASVSPSGGLDAGTNTGSSNSQ 1139

QY 389 SRTHPLOSPT--LVSGVAASPGCVFY-----BENGI 419  
Db 1140 DFPAPPGSAPMAALDGY-PGYFGSPYFGASGPQDYATAGQMPPSONNQTPIHGA 1198

QY 420 GGVAPPSSTSYLLPLE--AATGIPPSILNPHNQ--PFVNPDTPAI-----Y 466  
Db 1199 AAAVAAGDNISVSNPEEDPIAAGGPGS--GTGPGPGQGGPFAAGSAGAVGAVGGGPP 1257

QY 467 NPPTSQPIRSAM--VGSQOQPPQOQ--PSPOQOQVQ-----PPQ 505  
Db 1258 HPPPHSPHTAAQAAQGHQOQHPOHHPGLPQPPPOQOQOQOQPPPSVGGGPPAP 1317

QY 506 QWAGPLVTVSGQLQASSQSVQPAVS--FPP--OHLPLVS--PTQHFMRDQVATQFG 558  
Db 1318 QOHGP-----GVPPSPQOHVRFPAAGARYPPGSGSPYFVSKTSPSPSGAGAGQVQ 1371

QY 559 QWTLRS-----QSSGETPE-----PPSGP-----VYPSLSMPD-- 587  
Db 1372 SSDQVYATGPPPGQPFQGGPQYPPQNRNMYPPYGPGEGBAPPTGANGYGFSGSPYQPP 1431

QY 588 --AQPSYIATGQQLPTGSGSGSP-----PISQVLOPPSPSGFVQPPPAQM 637  
Db 1432 GGPQPTQTVAGGP--PAGGAPGAPSSGAYPTGRSQDQYQPPD-----OSPPRRH 1483

QY 638 PLYVYPSGQYF-----TSTTQYRPM----- 658  
Db 1484 PPFIKDSQYFPEGNARPOTYGAMQSGTQYRQYPSPPAPQWNGAPPGGAAPPGAPHG 1543

QY 659 APVQINA-----QSSQMPQAQAGVQPVUSGQGFGLGLGVQ 698  
Db 1544 PPIQPPAGVAGWDQHRYPQPGPPPPQOQOQOQOQOQOQPPYQV-AGPG-----QQ 1595

QY 699 PPOS-QNVINNOGTPTVQVWVSYPVMSYQVPMTOGSGS-----GLP-----QOSYQO--P1 747  
Db 1596 PPOAPPQAMQNNPGQTASGIA--PPGSLRPPSGGQGNRMGMQAQOQOQOQOQGVQ 1653

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OW protein - protein search, using sw model

Run on: July 5, 2006, 09:49:57 ; Search time 99 Seconds

(without alignments)  
3803.977 Million cell updates/sec

Title: US-10-509-950-1

Perfect score: 4235  
Sequence: 1 MSQGDNLNQAIAEBGGTGE.....SASCRITVCASMSNAGVOKF 813

Scoring table: BIOSOM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing of 45 summaries

Database :

1: Published Applications AA Main:  
2: /BMC\_Celerra\_SIDS3/prodata/2/pubppaa/US07\_PUBCOMB.pep.\*  
3: /BMC\_Celerra\_SIDS3/prodata/2/pubppaa/US08\_PUBCOMB.pep.\*  
4: /BMC\_Celerra\_SIDS3/prodata/2/pubppaa/US09\_PUBCOMB.pep.\*  
5: /BMC\_Celerra\_SIDS3/prodata/2/pubppaa/US10A\_PUBCOMB.pep.\*  
6: /BMC\_Celerra\_SIDS3/prodata/2/pubppaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description         |
|------------|--------|-------------|--------|-------|---------------------|
| 1          | 4235   | 100.0       | 813    | 4     | US-10-112-372-6     |
| 2          | 4235   | 100.0       | 813    | 5     | US-10-509-950-1     |
| 3          | 4208   | 99.4        | 847    | 4     | US-10-112-372-2     |
| 4          | 4111   | 97.1        | 793    | 4     | US-10-112-372-8     |
| 5          | 4084   | 96.4        | 827    | 4     | US-10-112-372-4     |
| 6          | 3421   | 80.8        | 807    | 4     | US-10-112-372-11    |
| 7          | 3414.5 | 80.6        | 651    | 4     | US-10-112-372-13    |
| 8          | 3398.5 | 80.2        | 648    | 4     | US-10-112-372-12    |
| 9          | 2513.5 | 59.4        | 645    | 5     | US-10-450-763-35346 |
| 10         | 2047   | 48.3        | 438    | 5     | US-10-450-763-32449 |
| 11         | 1831.5 | 43.2        | 1099   | 5     | US-10-756-149-5489  |
| 12         | 1493   | 35.3        | 976    | 5     | US-10-491-566-92    |
| 13         | 1493   | 35.3        | 976    | 5     | US-10-491-566-92    |
| 14         | 1455   | 34.4        | 1027   | 5     | US-10-490-318-10    |
| 15         | 1451   | 34.3        | 958    | 4     | US-10-144-194A-90   |
| 16         | 1451   | 34.3        | 958    | 4     | US-10-491-566-90    |
| 17         | 1420.5 | 33.5        | 433    | 5     | US-10-450-763-32447 |
| 18         | 1020   | 24.1        | 366    | 5     | US-10-450-763-35345 |
| 19         | 944    | 22.3        | 371    | 5     | US-10-450-763-32448 |
| 20         | 473    | 11.2        | 107    | 4     | US-10-104-047-3736  |
| 21         | 473    | 11.2        | 107    | 6     | US-11-072-512-3736  |
| 22         | 430    | 10.2        | 89     | 4     | US-10-112-372-10    |
| 23         | 430    | 10.2        | 89     | 4     | US-10-112-372-14    |
| 24         | 410.5  | 9.7         | 1805   | 6     | US-11-097-143-22578 |
| 25         | 346.5  | 8.2         | 94     | 5     | US-10-450-763-32444 |
| 26         | 346.5  | 8.2         | 94     | 5     | US-10-450-763-35343 |
| 27         | 344.5  | 8.1         | 88     | 4     | US-10-112-372-15    |

|    |       |     |      |   |                     |                   |
|----|-------|-----|------|---|---------------------|-------------------|
| 28 | 311.5 | 7.4 | 78   | 3 | US-09-864-408A-5512 | Sequence 5512, Ap |
| 29 | 308   | 7.3 | 79   | 5 | US-10-450-763-32445 | Sequence 32445, A |
| 30 | 308   | 7.3 | 79   | 5 | US-10-450-763-35344 | Sequence 35344, A |
| 31 | 293   | 6.9 | 2004 | 5 | US-10-469-469-250   | Sequence 250, App |
| 32 | 291   | 6.9 | 2703 | 6 | US-11-097-143-16689 | Sequence 16689, A |
| 33 | 286   | 6.8 | 1390 | 4 | US-10-092-900A-224  | Sequence 224, App |
| 34 | 274.5 | 6.5 | 1566 | 5 | US-10-491-545A-63   | Sequence 63, App1 |
| 35 | 268   | 6.3 | 3394 | 4 | US-10-408-765A-529  | Sequence 529, App |
| 36 | 265   | 6.3 | 1026 | 5 | US-10-415-656-2     | Sequence 2, App1  |
| 37 | 263.5 | 6.2 | 926  | 6 | US-11-097-143-22197 | Sequence 22197, A |
| 38 | 260   | 6.1 | 1250 | 4 | US-10-112-944-249   | Sequence 249, App |
| 39 | 259.5 | 6.1 | 1240 | 4 | US-10-369-993-4031  | Sequence 4031, Ap |
| 40 | 256.5 | 6.1 | 2759 | 5 | US-10-450-763-47507 | Sequence 47507, A |
| 41 | 255   | 6.0 | 1189 | 4 | US-10-408-765A-2272 | Sequence 2272, Ap |
| 42 | 251.5 | 5.9 | 2285 | 5 | US-10-773-446-101   | Sequence 101, App |
| 43 | 251   | 5.9 | 2280 | 6 | US-11-097-143-21742 | Sequence 11742, A |
| 44 | 249   | 5.9 | 3080 | 6 | US-11-097-143-21423 | Sequence 21423, A |
| 45 | 248   | 5.9 | 1790 | 6 | US-11-097-143-42009 | Sequence 42009, A |

## ALIGNMENTS

RESULT 1  
US-10-112-372-6  
; Sequence 6, Application US/10112372  
; Publication No. US20030186249A1  
; GENERAL INFORMATION:  
; APPLICANT: Origene Technologies, Inc.  
; FILE REFERENCE: Human TRAP Genes and Polypeptides  
; CURRENT APPLICATION NUMBER: US/10/112,372  
; CURRENT FILING DATE: 2002-04-01  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 813  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-112-372-6

Query Match 100.0%; Score 4235; DB 4; Length 813;  
Best Local Similarity 100.0%; Pred. No. 8.2e-234;  
Matches 813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |     |   |     |
|----|-----|---|-----|
| QY | 1   | 1ME0G0LQAIAEBGTE0ETATPENGIVKSESLDEEKLQRLLEAQNQERRKSSGA      | 60  |
| Db | 1   | 1ME0G0LQAIAEBGTE0ETATPENGIVKSESLDEEKLQRLLEAQNQERRKSSGA      | 60  |
| QY | 61  | GKGKLTSLAVCESSARPGESLQDESIIHQLSFFSSIOEDKSRKDSEREKDXN        | 120 |
| Db | 61  | GKGKLTSLAVCESSARPGESLQDESIIHQLSFFSSIOEDKSRKDSEREKDXN        | 120 |
| QY | 121 | KKTESEKPIKRMSLKDCSQEYTDSTGIDHAEFLINTLKNNSRDRMILKNOEIIIDFIAD | 180 |
| Db | 121 | KKTESEKPIKRMSLKDCSQEYTDSTGIDHAEFLINTLKNNSRDRMILKNOEIIIDFIAD | 180 |
| QY | 181 | NNNHKKKPPQMSXYRMLVHRYAAYFGLDHNDQCKSYIINKTSSTRIPPEORFCHLKD   | 240 |
| Db | 181 | NNNHKKKPPQMSXYRMLVHRYAAYFGLDHNDQCKSYIINKTSSTRIPPEORFCHLKD   | 240 |
| QY | 241 | EKGESQKRFILKRNSSIDKEDNSVCSQESILFENSRLLEDSNINETYKKRQLFRGN    | 300 |
| Db | 241 | EKGESQKRFILKRNSSIDKEDNSVCSQESILFENSRLLEDSNINETYKKRQLFRGN    | 300 |
| QY | 301 | RDGSGRTSGSRSSSENEIKMSDHPAWSTTSDSSNNRLKPAATKTASFGGITVLTTRGD  | 360 |
| Db | 301 | RDGSGRTSGSRSSSENEIKMSDHPAWSTTSDSSNNRLKPAATKTASFGGITVLTTRGD  | 360 |
| QY | 361 | STSSRSRTSGKSKASSESSSSAGSSGSLSRTHPLOSRTLVSGVAAGSGCVYPENNGIG  | 420 |
| Db | 361 | STSSRSRTSGKSKASSESSSSAGSSGSLSRTHPLOSRTLVSGVAAGSGCVYPENNGIG  | 420 |



QY 421 GQVAPSTSYILLPLEAATGIPPGSILLNPHTCOPFVNPDPGTPAIYNPPTSQOPLASAMV 480  
DB 421 GQVAPSTSYILLPLEAATGIPPGSILLNPHTCOPFVNPDPGTPAIYNPPTSQOPLASAMV 480  
QY 481 GQSQOQPPQOQSPPOQOQVQPPQPMAGPLVTQSVQGLQASSQSVOYPAVSFPPOHLLP 540  
DB 481 GQSQOQPPQOQSPPOQOQVQPPQPMAGPLVTQSVQGLQASSQSVOYPAVSFPPOHLLP 540  
QY 541 VSPTHCFPMKDDVAATQFGQMTLSRSSGGETPEPPSGPVYPSLMPQPAQOQPSYVIASTGQ 600  
DB 541 VSPTHCFPMKDDVAATQFGQMTLSRSSGGETPEPPSGPVYPSLMPQPAQOQPSYVIASTGQ 600  
QY 601 QLEPTGFGSSGPPISQOVLQPPSPPOGFPVQPPAPMPVYVYPSGQYPTSTTQOYRPMAP 660  
DB 601 QLEPTGFGSSGPPISQOVLQPPSPPOGFPVQPPAPMPVYVYPSGQYPTSTTQOYRPMAP 660  
QY 661 VOYNAGRSQOQMPQAAQOAGYQPVLSGQGFQGLIGVQOQPPQSONVINNOGCTPVQSVYMS 720  
DB 661 VOYNAGRSQOQMPQAAQOAGYQPVLSGQGFQGLIGVQOQPPQSONVINNOGCTPVQSVYMS 720  
QY 721 YPTMSSYQVPMTOGSGGLPQOQSTQOQIMLPNOAGQSLPATGMVYCNVTPPTPONNRL 780  
DB 721 YPTMSSYQVPMTOGSGGLPQOQSTQOQIMLPNOAGQSLPATGMVYCNVTPPTPONNRL 780  
QY 781 IGPCHCPSSTVPVMSASCRTNCASMSNAGQVXF 813  
DB 781 IGPCHCPSSTVPVMSASCRTNCASMSNAGQVXF 813

RESULT 2  
US-10-509-950-1

; Sequence 1, Application US/10509950  
; Publication No. US20060024305A1  
; GENERAL INFORMATION:  
; APPLICANT: Evotec Neurosciences GmbH  
; TITLE OF INVENTION: CAMP-Regulated Phosphoprotein for Diagnostic and  
; FILE REFERENCE: 020880ep  
; CURRENT APPLICATION NUMBER: US/10/509,950  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 813  
; TYPE: PRP  
; ORGANISM: Homo sapiens  
US-10-509-950-1

Query Match 100.0%; Score 4235; DB 5; Length 813;  
Best Local Similarity 100.0%; Pred. No. 8.2e-234;  
Matches 813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSEQGLNOLAIAEBGTEOETATPENGIVKSESLDEBEKLELORLEAQNQERRKSKSGA 60  
DB 1 MSEQGLNOLAIAEBGTEOETATPENGIVKSESLDEBEKLELORLEAQNQERRKSKSGA 60  
QY 61 GKGLTRSLAVCESSARPGESLQOQESIHLQSSFSLOEBDKSRKODSEKEKDXN 120  
DB 61 GKGLTRSLAVCESSARPGESLQOQESIHLQSSFSLOEBDKSRKODSEKEKDXN 120  
QY 121 KQTSSEKPKIRMLSKDCSOEYTDSTGIDLHEFLINTLKNSRDRMILLKXEOETIDFIAD 180  
DB 121 KQTSSEKPKIRMLSKDCSOEYTDSTGIDLHEFLINTLKNSRDRMILLKXEOETIDFIAD 180  
QY 181 NNNHYKKFPOMSSYQRMVHRVAAYFGLDHNVDQTKSVIINTTSTRIFEOQFCEHLKD 240  
DB 181 NNNHYKKFPOMSSYQRMVHRVAAYFGLDHNVDQTKSVIINTTSTRIFEOQFCEHLKD 240  
QY 241 EKSEESQKRFILKRDNSIDKEDNOSVCSQESLFEVNSRLLEDNSINICNETYKKRQLFRGN 300  
DB 241 EKSEESQKRFILKRDNSIDKEDNOSVCSQESLFEVNSRLLEDNSINICNETYKKRQLFRGN 300  
QY 301 RDGSGRTSGROSSSENELKMSDHQRAMSSTDSDSNRLKPMATKTASFGGITVLTTRGD 360

DB 301 RDGSGRTSGROSSSENELKMSDHQRAMSSTDSDSNRLKPMATKTASFGGITVLTTRGD 360  
QY 361 STSSTSTGKLSKAGSSSSSAGSSLSRTTHPPLQSTPLVSGVAAGSPCCVYPPENGIG 420  
DB 361 STSSTSTGKLSKAGSSSSSAGSSLSRTTHPPLQSTPLVSGVAAGSPCCVYPPENGIG 420  
QY 421 GQVAPSTSYILLPLEAATGIPPGSILLNPHTCOPFVNPDPGTPAIYNPPTSQOPLASAMV 480  
DB 421 GQVAPSTSYILLPLEAATGIPPGSILLNPHTCOPFVNPDPGTPAIYNPPTSQOPLASAMV 480  
QY 481 GQSQOQPPQOQSPPOQOQVQPPQPMAGPLVTQSVQGLQASSQSVOYPAVSFPPOHLLP 540  
DB 481 GQSQOQPPQOQSPPOQOQVQPPQPMAGPLVTQSVQGLQASSQSVOYPAVSFPPOHLLP 540  
QY 541 VSPTHCFPMKDDVAATQFGQMTLSRSSGGETPEPPSGPVYPSLMPQPAQOQPSYVIASTGQ 600  
DB 541 VSPTHCFPMKDDVAATQFGQMTLSRSSGGETPEPPSGPVYPSLMPQPAQOQPSYVIASTGQ 600  
QY 601 QLEPTGFGSSGPPISQOVLQPPSPPOGFPVQPPAPMPVYVYPSGQYPTSTTQOYRPMAP 660  
DB 601 QLEPTGFGSSGPPISQOVLQPPSPPOGFPVQPPAPMPVYVYPSGQYPTSTTQOYRPMAP 660  
QY 661 VOYNAGRSQOQMPQAAQOAGYQPVLSGQGFQGLIGVQOQPPQSONVINNOGCTPVQSVYMS 720  
DB 661 VOYNAGRSQOQMPQAAQOAGYQPVLSGQGFQGLIGVQOQPPQSONVINNOGCTPVQSVYMS 720  
QY 721 YPTMSSYQVPMTOGSGGLPQOQSTQOQIMLPNOAGQSLPATGMVYCNVTPPTPONNRL 780  
DB 721 YPTMSSYQVPMTOGSGGLPQOQSTQOQIMLPNOAGQSLPATGMVYCNVTPPTPONNRL 780  
QY 781 IGPCHCPSSTVPVMSASCRTNCASMSNAGQVXF 813  
DB 781 IGPCHCPSSTVPVMSASCRTNCASMSNAGQVXF 813

RESULT 3

US-10-112-372-2  
; Sequence 2, Application US/10112372  
; Publication No. US20030186249A1  
; GENERAL INFORMATION:  
; APPLICANT: Origene Technologies, Inc.  
; TITLE OF INVENTION: Human Tarpp Genes and Polypeptides  
; FILE REFERENCE: 16U 105 R1  
; CURRENT APPLICATION NUMBER: US/10/112,372  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 847  
; TYPE: PRP  
; ORGANISM: Homo sapiens  
US-10-112-372-2

Query Match 99.4%; Score 4208; DB 4; Length 847;  
Best Local Similarity 96.0%; Pred. No. 3e-232;  
Matches 813; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 1 MSEQGLNOLAIAEBGTEOETATPENGIVKSESLDEBEKLELORLEAQNQERRKSKSGA 60  
DB 1 MSEQGLNOLAIAEBGTEOETATPENGIVKSESLDEBEKLELORLEAQNQERRKSKSGA 60  
QY 61 GKGLTRSLAVCESSARPGESLQOQESIHLQSSFSLOEBDKSRKODSEKEKDXN 120  
DB 61 GKGLTRSLAVCESSARPGESLQOQESIHLQSSFSLOEBDKSRKODSEKEKDXN 120  
QY 121 KQTSSEKPKIRMLSKDCSOEYTDSTGIDLHEFLINTLKNSRDRMILLKXEOETIDFIAD 180  
DB 121 KQTSSEKPKIRMLSKDCSOEYTDSTGIDLHEFLINTLKNSRDRMILLKXEOETIDFIAD 180  
QY 181 NNNHYKKFPOMSSYQRMVHRVAAYFGLDHNVDQTKSVIINTTSTRIFEOQFCEHLKD 240  
DB 181 NNNHYKKFPOMSSYQRMVHRVAAYFGLDHNVDQTKSVIINTTSTRIFEOQFCEHLKD 240



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2006, 09:44:02 ; Search time 3004 Seconds

(Without alignments)  
9988.817 Million cell updates/sec

Title: US-10-509-950-2

Perfect score: 2441.6

Sequence: 1 atgtctgagcaggaagacc.....gttcgagctcaattctga 2442

Scoring table: IDENTITY NUC

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Published Applications NA Main:

1: /EMC\_Celerra\_SIDS3/prodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /EMC\_Celerra\_SIDS3/prodata/2/pubpna/US09\_PUBCOMB.seq.\*  
3: /EMC\_Celerra\_SIDS3/prodata/2/pubpna/US09\_PUBCOMB.seq.\*  
4: /EMC\_Celerra\_SIDS3/prodata/2/pubpna/US09\_PUBCOMB.seq.\*  
5: /EMC\_Celerra\_SIDS3/prodata/2/pubpna/US09\_PUBCOMB.seq.\*  
6: /EMC\_Celerra\_SIDS3/prodata/2/pubpna/US10\_PUBCOMB.seq.\*  
7: /EMC\_Celerra\_SIDS3/prodata/2/pubpna/US10\_PUBCOMB.seq.\*  
8: /EMC\_Celerra\_SIDS3/prodata/2/pubpna/US10\_PUBCOMB.seq.\*  
9: /EMC\_Celerra\_SIDS3/prodata/2/pubpna/US10\_PUBCOMB.seq.\*  
10: /EMC\_Celerra\_SIDS3/prodata/2/pubpna/US10\_PUBCOMB.seq.\*  
11: /EMC\_Celerra\_SIDS3/prodata/2/pubpna/US10\_PUBCOMB.seq.\*  
12: /EMC\_Celerra\_SIDS3/prodata/2/pubpna/US10\_PUBCOMB.seq.\*  
13: /EMC\_Celerra\_SIDS3/prodata/2/pubpna/US10\_PUBCOMB.seq.\*  
14: /EMC\_Celerra\_SIDS3/prodata/2/pubpna/US10\_PUBCOMB.seq.\*  
15: /EMC\_Celerra\_SIDS3/prodata/2/pubpna/US10\_PUBCOMB.seq.\*  
16: /EMC\_Celerra\_SIDS3/prodata/2/pubpna/US10\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Length | DB ID | Description        |
|------------|--------|--------------|-------|--------------------|
| 1          | 2441.6 | 100.0        | 11    | US-10-509-950-2    |
| 2          | 2441.6 | 100.0        | 11    | US-10-509-950-3    |
| 3          | 2441.2 | 100.0        | 7     | US-10-112-372-5    |
| 4          | 2334.8 | 95.2         | 7     | US-10-112-372-1    |
| 5          | 2311.2 | 94.7         | 7     | US-10-112-372-7    |
| 6          | 2287.2 | 93.7         | 7     | US-10-112-372-3    |
| 7          | 1746.4 | 71.5         | 2548  | US-10-450-763-2081 |
| 8          | 1366.2 | 56.0         | 1942  | US-10-450-763-4978 |
| 9          | 750.2  | 30.7         | 1305  | US-10-450-763-2079 |
| 10         | 650.4  | 26.6         | 1488  | US-10-450-763-2080 |
| 11         | 650.4  | 26.6         | 1488  | US-10-450-763-4977 |
| 12         | 362.8  | 14.9         | 4666  | US-10-144-194A-91  |
| 13         | 308.8  | 12.6         | 4330  | US-10-491-566-91   |
| 14         | 274.8  | 11.3         | 3829  | US-10-490-318-45   |
| 15         | 262.2  | 10.7         | 2014  | US-10-104-047-1766 |
| 16         | 262.2  | 10.7         | 2014  | US-11-072-512-1766 |

|    |       |      |      |    |                       |                    |
|----|-------|------|------|----|-----------------------|--------------------|
| 18 | 260.6 | 10.7 | 1006 | 7  | US-10-112-372-9       | Sequence 9, Appli  |
| 19 | 260.6 | 10.7 | 1048 | 9  | US-10-335-053-17      | Sequence 17, Appli |
| 20 | 260.6 | 10.7 | 4276 | 7  | US-10-144-194A-89     | Sequence 89, Appli |
| 21 | 260.6 | 10.7 | 4276 | 9  | US-10-491-566-89      | Sequence 89, Appli |
| 22 | 257.4 | 10.5 | 371  | 10 | US-10-450-763-2077    | Sequence 2077, Ap  |
| 23 | 257.4 | 10.5 | 371  | 10 | US-10-450-763-4976    | Sequence 4976, Ap  |
| 24 | 233.8 | 9.6  | 579  | 7  | US-10-029-386-4603    | Sequence 4603, Ap  |
| 25 | 231.2 | 9.5  | 274  | 7  | US-10-029-386-18359   | Sequence 18359, A  |
| 26 | 229.4 | 9.4  | 567  | 10 | US-10-450-763-2078    | Sequence 2078, Ap  |
| 27 | 207.2 | 8.5  | 308  | 10 | US-10-450-763-2076    | Sequence 2076, Ap  |
| 28 | 207.2 | 8.5  | 308  | 10 | US-10-450-763-4975    | Sequence 4975, Ap  |
| 29 | 203.2 | 8.3  | 1305 | 10 | US-10-450-763-2079    | Sequence 2079, Ap  |
| 30 | 203.2 | 8.3  | 1942 | 10 | US-10-450-763-4978    | Sequence 4978, Ap  |
| 31 | 177   | 7.2  | 404  | 3  | US-09-918-995-34085   | Sequence 34085, A  |
| 32 | 160   | 6.6  | 1030 | 16 | US-11-136-527-3103    | Sequence 3103, Ap  |
| 33 | 158.4 | 6.5  | 1195 | 16 | US-11-136-527-3103    | Sequence 3103, Ap  |
| 34 | 150.8 | 6.2  | 285  | 3  | US-09-864-408A-5511   | Sequence 5511, Ap  |
| 35 | 150.4 | 6.2  | 1641 | 4  | US-09-925-065A-721030 | Sequence 721030, A |
| 36 | 150.4 | 6.2  | 1641 | 5  | US-09-925-065A-721030 | Sequence 721030, A |
| 37 | 144.8 | 5.5  | 510  | 7  | US-10-029-386-4818    | Sequence 4818, Ap  |
| 38 | 129   | 5.3  | 302  | 7  | US-10-029-386-18574   | Sequence 18574, A  |
| 39 | 119.2 | 4.9  | 398  | 8  | US-10-242-535A-21191  | Sequence 21191, A  |
| 40 | 119.2 | 4.9  | 398  | 8  | US-10-085-783A-21191  | Sequence 21191, A  |
| 41 | 115.4 | 4.7  | 569  | 6  | US-10-040-739-1301    | Sequence 301, App  |
| 42 | 96.4  | 3.9  | 581  | 12 | US-10-301-480-52544   | Sequence 52544, A  |
| 43 | 96.4  | 3.9  | 581  | 12 | US-10-301-480-665953  | Sequence 665953, A |
| 44 | 96.4  | 3.9  | 624  | 6  | US-10-027-632-198151  | Sequence 198151, A |
| 45 | 96.4  | 3.9  | 624  | 7  | US-10-027-632-198151  | Sequence 198151, A |

## ALIGNMENTS

|  |     |   |     |  |  |  |  |  |  |
|--|-----|---|-----|--|--|--|--|--|--|
| RESULT 1   |     |   |     |  |  |  |  |  |  |
| US-10-509-950-2  |     |   |     |  |  |  |  |  |  |
| Sequence 2, Application US/10509950                                  |     |   |     |  |  |  |  |  |  |
| Publication No. US20060024305A1                                      |     |   |     |  |  |  |  |  |  |
| GENERAL INFORMATION:   |     |   |     |  |  |  |  |  |  |
| APPLICANT: Evotec NeuroSciences GmbH                                 |     |   |     |  |  |  |  |  |  |
| TITLE OF INVENTION: CAMP-Regulated Phosphoprotein for Diagnostic and |     |   |     |  |  |  |  |  |  |
| TITLE OF INVENTION: Therapeutic Use in Neurodegenerative Diseases    |     |   |     |  |  |  |  |  |  |
| FILE REFERENCE: 020880ep   |     |   |     |  |  |  |  |  |  |
| CURRENT APPLICATION NUMBER: US/10/509,950                            |     |   |     |  |  |  |  |  |  |
| CURRENT FILING DATE: 2004-10-04                                      |     |   |     |  |  |  |  |  |  |
| NUMBER OF SEQ ID NOS: 16   |     |   |     |  |  |  |  |  |  |
| SOFTWARE: PatentIn Ver. 2.1  |     |   |     |  |  |  |  |  |  |
| SEQ ID NO 2  |     |   |     |  |  |  |  |  |  |
| LENGTH: 2442   |     |   |     |  |  |  |  |  |  |
| TYPE: DNA  |     |   |     |  |  |  |  |  |  |
| ORGANISM: Homo sapiens   |     |   |     |  |  |  |  |  |  |
| US-10-509-950-2  |     |   |     |  |  |  |  |  |  |
| Query Match  |     |   |     |  |  |  |  |  |  |
| Best Local Similarity 100.0%; Score 2441.6; DB 11; Length 2442;      |     |   |     |  |  |  |  |  |  |
| Matches 2442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;        |     |   |     |  |  |  |  |  |  |
| Oy   | 1   | ATGCTGAGCAGGAGACTGTAATCAGCAATACAGAGAGGAGCTGAGCAGAG    | 60  |  |  |  |  |  |  |
| Db   | 1   | ATGCTGAGCAGGAGACTGTAATCAGCAATACAGAGAGGAGCTGAGCAGAG    | 60  |  |  |  |  |  |  |
| Oy   | 61  | ACGGCACTCAGAGAGCGCATTTGTAATCAGAAAGCTGATGAAGAGGAACTG   | 120 |  |  |  |  |  |  |
| Db   | 61  | ACGGCACTCAGAGAGCGCATTTGTAATCAGAAAGCTGATGAAGAGGAACTG   | 120 |  |  |  |  |  |  |
| Oy   | 121 | GAAGTGAAGAGGCGCTGAGGCTCAGAAATCAAGAAAGAAATCAATCAGAGACA | 180 |  |  |  |  |  |  |
| Db   | 121 | GAAGTGAAGAGGCGCTGAGGCTCAGAAATCAAGAAAGAAATCAATCAGAGACA | 180 |  |  |  |  |  |  |
| Oy   | 181 | GGAAGAGTAACTGATCCAGCTGCTGCTGTGAGAGATCTTCCAGACAGAGA    | 240 |  |  |  |  |  |  |
| Db   | 181 | GGAAGAGTAACTGATCCAGCTGCTGCTGTGAGAGATCTTCCAGACAGAGA    | 240 |  |  |  |  |  |  |
| Oy   | 241 | GGTGAAGTCTTCAGATCAGAAATCAATTCATTACAGCTTTCCAGAGCTG     | 300 |  |  |  |  |  |  |

|    |      |   |      |
|----|------|---|------|
| Db | 241  | GGTAAAGCTCTTCAGGATCAGGAATCAAATTCATTACGCTTCCAGTTTCCAGCTG       | 300  |
| Qy | 301  | CAAGAGGAGTAATCTGAGGAAGATGACTCTGAAAGAGAAAAAGAAAGATAAAAAC       | 360  |
| Db | 301  | CAAGAGGAGTAATCTGAGGAAGATGACTCTGAAAGAGAAAAAGAAAGATAAAAAC       | 360  |
| Qy | 361  | AAAGATAAACCTCTGAAAAAACCCAGATCGAATGTTATCAAAAGATTCGACCGAAGAA    | 420  |
| Db | 361  | AAAGATAAACCTCTGAAAAAACCCAGATCGAATGTTATCAAAAGATTCGACCGAAGAA    | 420  |
| Qy | 421  | TACACGATTTACAGGATAGACTTTCACAGGTTTCGATTAACAATTAAAGAAATAT       | 480  |
| Db | 421  | TACACGATTTACAGGATAGACTTTCACAGGTTTCGATTAACAATTAAAGAAATAT       | 480  |
| Qy | 481  | TCCAGGGACAGATGATPACTTTTGAAAAATGAGCAGGAAATATATGATTCATGTCTGAC   | 540  |
| Db | 481  | TCCAGGGACAGATGATPACTTTTGAAAAATGAGCAGGAAATATATGATTCATGTCTGAC   | 540  |
| Qy | 541  | AACATATATCATTTATPAAAAAGTCCCTCAGATGTCATCGATCAGAGATGCTGTCCAT    | 600  |
| Db | 541  | AACATATATCATTTATPAAAAAGTCCCTCAGATGTCATCGATCAGAGATGCTGTCCAT    | 600  |
| Qy | 601  | CGAGTGGCAGCTTATTTTGGATTGGATCACAATGTGGATCAAAACAGAAAAATCGTTATC  | 660  |
| Db | 601  | CGAGTGGCAGCTTATTTTGGATTGGATCACAATGTGGATCAAAACAGAAAAATCGTTATC  | 660  |
| Qy | 661  | ATCAACAAGACACAGACACAGAAATACACAGGAAAGGTTTGTGAACATTTAAAAAGAT    | 720  |
| Db | 661  | ATCAACAAGACACAGACACAGAAATACACAGGAAAGGTTTGTGAACATTTAAAAAGAT    | 720  |
| Qy | 721  | GAAAAAGGTGAAGATCCACAGACGGTTTATCTTGAAACGAGATPAACTCTAGATTGAT    | 780  |
| Db | 721  | GAAAAAGGTGAAGATCCACAGACGGTTTATCTTGAAACGAGATPAACTCTAGATTGAT    | 780  |
| Qy | 781  | AAAGAAACAATTCAGTCAAGTTGCTCCACAGAAAGCCTTTTGTGAAAAACATGAGCTC    | 840  |
| Db | 781  | AAAGAAACAATTCAGTCAAGTTGCTCCACAGAAAGCCTTTTGTGAAAAACATGAGCTC    | 840  |
| Qy | 841  | TTGGAAGACAATAATATGCAATGAGACCTATTAAGAAAAACAGACTCTTCCGGGGAC     | 900  |
| Db | 841  | TTGGAAGACAATAATATGCAATGAGACCTATTAAGAAAAACAGACTCTTCCGGGGAC     | 900  |
| Qy | 901  | AGAGATGAGCTCAGGGAGAAACATCTGGAGTGCAGACAGACACTCAGAAAAATGAATCAAG | 960  |
| Db | 901  | AGAGATGAGCTCAGGGAGAAACATCTGGAGTGCAGACAGACACTCAGAAAAATGAATCAAG | 960  |
| Qy | 961  | TGATCTGACCAACAAAGGCTGTGAGCAGACAGACTTCGACAGTTCGAACCGCAATCTA    | 1020 |
| Db | 961  | TGATCTGACCAACAAAGGCTGTGAGCAGACAGACTTCGACAGTTCGAACCGCAATCTA    | 1020 |
| Qy | 1021 | AAAGCCCGCATGACCAAGACGGCAAGTTTGGGGCATCAGGTGTCTGACCAAGGGGTGAC   | 1080 |
| Db | 1021 | AAAGCCCGCATGACCAAGACGGCAAGTTTGGGGCATCAGGTGTCTGACCAAGGGGTGAC   | 1080 |
| Qy | 1081 | AGCATTCCAGTACTAGGATGATACGGGAAAGCTGTCCAAAGCAGGTTCCGATCTTCCAGC  | 1140 |
| Db | 1081 | AGCATTCCAGTACTAGGATGATACGGGAAAGCTGTCCAAAGCAGGTTCCGATCTTCCAGC  | 1140 |
| Qy | 1141 | AGTGCAGGCTCCTCAGATCGCTGTCCCGCACCCATCCACTCTCCAGAGCACCCCTA      | 1200 |
| Db | 1141 | AGTGCAGGCTCCTCAGATCGCTGTCCCGCACCCATCCACTCTCCAGAGCACCCCTA      | 1200 |
| Qy | 1201 | GTCTCAGGTGTGGCAGCTGTGCTTCCAGGCTGTGTGCTTATCCAGAGAAATGGAATAGG   | 1260 |
| Db | 1201 | GTCTCAGGTGTGGCAGCTGTGCTTCCAGGCTGTGTGCTTATCCAGAGAAATGGAATAGG   | 1260 |
| Qy | 1261 | GGCACAGTTGTCCACAGCACACAGCTACATCTCTTCCACTTGAAGCTGCACACGCG      | 1320 |
| Db | 1261 | GGCACAGTTGTCCACAGCACACAGCTACATCTCTTCCACTTGAAGCTGCACACGCG      | 1320 |
| Qy | 1321 | ATCCGCGCTGGAAGCATCTTCTTATCAACAACAGGCGACGCTTTGTGATCCGAT        | 1380 |
| Db | 1321 | ATCCGCGCTGGAAGCATCTTCTTATCAACAACAGGCGACGCTTTGTGATCCGAT        | 1380 |

|    |      |  |      |
|----|------|--|------|
| Db | 1231 | ATCCCGCCTGGAAAGCATCTCTTTATATCCACACACAGGCGAGCCCTTTGGAAATCCCGAT  | 1380 |
| Oy | 1381 | GGAACTCTCTGCAATATACAACTCCACCACAGTCAAGACCCCTTGGAAAGCCCATGTG     | 1440 |
| Db | 1381 | GGAACTCTCTGCAATATATACAACTCCACCACAGTCAAGACCCCTTGGAAAGCCCATGTG   | 1440 |
| Oy | 1441 | GGGCAAGTCCCAACAGCAGCCGCGACAGACAGAGCCCTCCCGGAGCCCAACAGCAGTTC    | 1500 |
| Db | 1441 | GGGCAAGTCCCAACAGCAGCCGCGACAGACAGAGCCCTCCCGGAGCCCAACAGCAGTTC    | 1500 |
| Oy | 1501 | CAGCCACCGCAGCCACAGATGGCAGAGCCCTCTGATCACTAGTCTGTCCAGGGGCTGAG    | 1560 |
| Db | 1501 | CAGCCACCGCAGCCACAGATGGCAGAGCCCTCTGATCACTAGTCTGTCCAGGGGCTGAG    | 1560 |
| Oy | 1561 | GCTTCTCTCCAGTCAAGTCAATATCCGCGAGTCTCTTCTCCACAGCACTTCACT         | 1620 |
| Db | 1561 | GCTTCTCTCCAGTCAAGTCAATATCCGCGAGTCTCTTCTCTCCACAGCACTTCACT       | 1620 |
| Oy | 1621 | GTTGTTCCAAAGCAGACATTTTCCCATGAGATGATGATGGCAACACAGTTTGGCGAATG    | 1680 |
| Db | 1621 | GTTGTTCCAAAGCAGACATTTTCCCATGAGATGATGATGGCAACACAGTTTGGCGAATG    | 1680 |
| Oy | 1681 | ACCCTGAGCCGGCAGTCTCTCGGGGAGAGTCTCTGAAACCCCATCAGGTCTGTCTACCA    | 1740 |
| Db | 1681 | ACCCTGAGCCGGCAGTCTCTCGGGGAGAGTCTCTGAAACCCCATCAGGTCTGTCTACCA    | 1740 |
| Oy | 1741 | TCCTCTCTTATGCAACAGCCGCGCCAGCAGACCCAGCTATGTATATGCTCTACAGGCCAG   | 1800 |
| Db | 1741 | TCCTCTCTTATGCAACAGCCGCGCCAGCAGACCCAGCTATGTATATGCTCTACAGGCCAG   | 1800 |
| Oy | 1801 | CAGCTTCTCTACAGAGGATTTCTCAGGCTGTGGCCCTTCCCATCTCCACAGCAGTCTCCAG  | 1860 |
| Db | 1801 | CAGCTTCTCTACAGAGGATTTCTCAGGCTGTGGCCCTTCCCATCTCCACAGCAGTCTCCAG  | 1860 |
| Oy | 1861 | CCCCCTCCCTACCAAGGGATTCGCGCAACAGCCTCCGCTGACACAGATGCTGATAT       | 1920 |
| Db | 1861 | CCCCCTCCCTACCAAGGGATTCGCGCAACAGCCTCCGCTGACACAGATGCTGATAT       | 1920 |
| Oy | 1921 | TATTATCCCATCTGTCAGTACCTTACCTTCACCAACAGCAACAGTACCGGCCCATGAGCCCG | 1980 |
| Db | 1921 | TATTATCCCATCTGTCAGTACCTTACCTTCACCAACAGCAACAGTACCGGCCCATGAGCCCG | 1980 |
| Oy | 1981 | GTTCACTACACAGCTTCAAGAGAGTCAACAGATGCCACAGGACACACAGCAAGCAGGTTAC  | 2040 |
| Db | 1981 | GTTCACTACACAGCTTCAAGAGAGTCAACAGATGCCACAGGACACACAGCAAGCAGGTTAC  | 2040 |
| Oy | 2041 | CAGCCAGTCTTGTCTGTGCTCAACAGGGATTTCAAGGCTTAATGAGATGCAAGACCACT    | 2100 |
| Db | 2041 | CAGCCAGTCTTGTCTGTGCTCAACAGGGATTTCAAGGCTTAATGAGATGCAAGACCACT    | 2100 |
| Oy | 2101 | CAGAGTCAAGACGTATATAATATAACCAACAGGAACCTCCGGTGCAGAAAGCTGATGTTCC  | 2160 |
| Db | 2101 | CAGAGTCAAGACGTATATAATATAACCAACAGGAACCTCCGGTGCAGAAAGCTGATGTTCC  | 2160 |
| Oy | 2161 | TACCCAAACAATGCTTCTTATCAGGTGCCAATGACCAAGGTTCTCAAGCATGCCCCAG     | 2220 |
| Db | 2161 | TACCCAAACAATGCTTCTTATCAGGTGCCAATGACCAAGGTTCTCAAGCATGCCCCAG     | 2220 |
| Oy | 2221 | CAGTATATACCAACAGCCAACTCATGCTACTTAACAGGCAAGTCAAGGCTCACTCCAGCC   | 2280 |
| Db | 2221 | CAGTATATACCAACAGCCAACTCATGCTACTTAACAGGCAAGTCAAGGCTCACTCCAGCC   | 2280 |
| Oy | 2281 | ACTGGAATGCGTGTATTACTGTATATGCAACCGCCACCCCTCAGAACCAACTTAAAGCTG   | 2340 |
| Db | 2281 | ACTGGAATGCGTGTATTACTGTATATGCAACCGCCACCCCTCAGAACCAACTTAAAGCTG   | 2340 |
| Oy | 2341 | ATTGGCCCAACTGCCCCCTTCCAGCACTGTCTCCAGTATGTACAGTACACTGCAGAACAAAC | 2400 |
| Db | 2341 | ATTGGCCCAACTGCCCCCTTCCAGCACTGTCTCCAGTATGTACAGTACACTGCAGAACAAAC | 2400 |
| Oy | 2401 | TGTGCAAGTATAGCAATGCTGTTGGCAAGGTCAAAATTTCTGA 2442               |      |
| Db | 2401 | TGTGCAAGTATAGCAATGCTGTTGGCAAGGTCAAAATTTCTGA 2442               |      |